

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:25:12 ; Search time 58.0357 Seconds
(without alignments)
888.869 Million cell updates/sec

Title: US-09-805-427a-2

Perfect score: 1745

Sequence: 1 MTDSVKIRAWGRRLMIGTA.....YWGQALNAMKGLQSLGAG 325

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1078588

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database : A Geneseq_19Jun03:*

```

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1745	100.0	325	18	AAW18164
2	1745	100.0	325	19	AAW63032
3	1745	100.0	325	19	AAW46776
4	1745	100.0	325	20	AAW14853
5	1745	100.0	325	22	AAW47556
6	1745	100.0	325	23	ABW73459
7	1745	100.0	325	23	AAW50730
8	1745	100.0	325	24	ABW56406
9	1706	97.8	323	23	ABW73461

10	1699	97.4	323	20	AAW14855	Antigen 85B protei
11	1602	91.8	325	12	AAW10486	Tuberculin active
12	1595	91.4	325	11	AAW08099	Mycobacterium deri
13	1550	88.8	403	19	AAW72943	Mycobacterium tube
14	1550	88.8	403	20	AAW21963	Amino acid sequenc
15	1550	88.8	404	19	AAW72942	Mycobacterium tube
16	1550	88.8	404	20	AAW21962	Amino acid sequenc
17	1522	87.2	330	19	AAW46777	Tuberculin active
18	1510	86.5	283	10	AAW90109	Antigen 85B protei
19	1454	83.3	327	20	AAW14851	M leprae 85B prote
20	1454	83.3	327	23	ABW73457	Mycobacterium-orig
21	1437	82.3	285	23	ABW83662	Mycobacterium tube
22	1407	80.6	337	18	ABW56407	Mycobacterium tube
23	1407	80.6	338	24	AAW18165	Mycobacterium tube
24	1407	80.6	338	19	AAW63033	Mycobacterium tube
25	1407	80.6	338	20	AAW14852	Antigen 85A protei
26	1407	80.6	338	20	AAW14854	Antigen 85A protei
27	1407	80.6	338	22	ABW47555	Ag85A. Mycobacter
28	1407	80.6	338	22	ABW82789	Mycobacterium tube
29	1407	80.6	338	23	ABW73458	M tuberculosis 85A
30	1407	80.6	338	23	ABW73460	M bovis 85A protei
31	1407	80.6	338	23	AAW50729	Mycobacterium tube
32	1401	80.3	338	12	AAW11296	Recombinant M tube
33	1362.5	78.1	353	12	AAW11295	Recombinant M tube
34	1335.5	76.5	330	20	AAW14850	Antigen 85A protei
35	1335.5	76.5	330	23	AAW73456	Mycobacterium tube
36	1323	75.8	295	23	AAW50759	M tuberculosis an
37	1322	75.8	302	21	AAW96926	M. tuberculosis an
38	1198	68.7	340	22	ABW47557	Ag85C. Mycobacter
39	1198	68.7	340	23	AAW50731	Mycobacterium tube
40	1195	68.5	340	13	AAW26167	85-C. Mycobacteri
41	1195	68.5	340	20	AAW14857	Antigen 85C protei
42	1195	68.5	340	23	ABW73463	M tuberculosis 85C
43	1179	67.6	326	19	AAW60112	Mycobacterium vacc
44	1179	67.6	326	20	AAW14858	M. vaccae antigen
45	1179	67.6	326	23	ABW73464	M vaccae antigen 8
46	1170	67.0	333	20	AAW14856	Antigen 85C protei
47	1170	67.0	333	23	ABW73462	M leprae 85C prote
48	1151.5	66.0	334	19	AAW60114	Mycobacterium vacc
49	1151.5	66.0	334	20	AAW14860	M. vaccae antigen
50	1151.5	66.0	334	23	ABW73466	M. vaccae antigen 8
51	644	36.9	161	19	AAW60113	Mycobacterium vacc
52	644	36.9	161	20	AAW14859	M. vaccae antigen
53	644	36.9	161	23	ABW73465	M vaccae antigen 8
54	637	36.5	161	19	AAW60127	Mycobacterium vacc
55	637	36.5	161	23	ABW73479	M vaccae GV-4P par
56	636	36.4	161	20	AAW14908	Amino acid sequenc
57	636	36.4	161	23	ABW73514	M vaccae GV-5P pro
58	635	36.4	160	20	AAW14873	Amino acid sequenc
59	519	29.7	239	18	AAW18162	Mycobacterium tube
60	519	29.7	239	19	AAW72899	Mycobacterium tube
61	519	29.7	239	19	AAW63037	Mycobacterium tube
62	519	29.7	239	20	AAW21916	Amino acid sequenc
63	474	27.2	657	14	AAW32657	PS1 protein from C
64	470	26.9	657	22	AAW92916	C glutamicum prote
65	415	23.8	341	22	AAW92919	C glutamicum prote
66	397	22.8	114	23	ABP03110	Human ORFX protein
67	303.5	17.4	411	22	AAW90895	C glutamicum prote
68	279	16.0	462	22	AAW90774	C glutamicum prote
69	277	15.9	483	22	AAW92137	C glutamicum prote
70	263	15.1	358	19	AAW56155	AA cell surface la
71	262	15.0	626	20	AAW98010	BCG Ag85B antigen-
72	262	15.0	626	24	AAW08321	M. bovis Ag85B-L.
73	238.5	13.7	365	22	AAW90128	C glutamicum prote
74	209	12.0	40	16	AAW85683	30 kD M. tuberculosis
75	209	12.0	40	17	AAW85687	Extended N-termina
76	209	12.0	40	16	AAW92883	Mycobacterium 30 k
77	209	12.0	40	18	AAW18191	N-terminal sequenc
78	209	12.0	40	19	AAW75579	M. tuberculosis 30
79	209	12.0	40	19	AAW75567	M. tuberculosis 30
80	209	12.0	40	24	ABW74416	M. tuberculosis ex
81	209	12.0	40	24	ABW74424	M. tuberculosis ex
82	209	12.0	40	24	ABW56331	M. tuberculosis ex

83	209	12.0	40	24	ABU56332	M. tuberculosis ex
84	209	12.0	40	24	ABU56343	M. tuberculosis ex
85	203	11.6	39	22	AAE05566	N-terminal amino a
86	202	11.6	40	18	AAW18192	N-terminal sequenc
87	202	11.6	40	24	ABU56344	M. tuberculosis ex
88	201	11.5	40	19	AAW75569	M. tuberculosis 32
89	198	11.3	40	19	AAW75568	M. tuberculosis 32
90	197	11.3	40	24	ABG74433	M. tuberculosis ex

ALIGNMENTS

RESULT 1	
AAW18164	
ID AAW18164	standard; Protein; 325 AA.
XX	
XX AC	AAW18164;
XX	
XX	13-AUG-1997 (first entry)
DT	
XX	
DE	Mycobacterium tuberculosis extracellular 30KD protein.
XX	
XX	Vaccine; vaccinating agent; M.tuberculosis; pathogen; bacteria;
KW KW	virus; fungus; protozoan; HIV.
XX	
XX	
OS	Mycobacterium tuberculosis.
XX	
FH	Key
FT	Location/Qualifiers
PT	41..325
FT	/label= Mature
XX	
XX	WO9637219-A1.
PN	
XX	
PD	28-NOV-1996.
XX	
XX	23-MAY-1996; 96WO-US07781.
XX	
PR	06-DEC-1995; 95US-0568357.
PR	23-MAY-1995; 95US-0447398.
PR	20-OCT-1995; 95US-0545926.
PR	31-OCT-1995; 95US-0551149.
XX	
PA	(REGC) UNIV CALIFORNIA.
XX	
PI	Harth G, Horwitz MA;
XX	
DR	WPI; 1997-020936/02.
DR	N-PSDB; AAW1598.
XX	
PT	Vaccines derived from M.tuberculosis major abundant extracellular
PT	proteins - are easy to prepare and less toxic than conventional
PT	killed or attenuated vaccines, useful for protecting against or
PT	treating Mycobacterial infections
XX	
PS	Claim 1; Page 35-36; 193pp; English.

A vaccinating agent for promoting an immune response in a mammal against Mycobacterium pathogens comprises at least one majorly abundant extracellular protein, i.e. the M.tuberculosis 110, 80, 71, 58, 45, 32A, 32B, 30, 24, 23.5, 23, 16, 14 or 12 kD proteins, or their analogues, homologues and subunits. The present sequence represents the 30 kD protein. The vaccinating agents are used to protect against (or to treat existing) infections by Mycobacterium (especially M. tuberculosis) while the epitopes can also be used to detect presence of an immune response to a Mycobacterium pathogen. The vectors, containing the DNA for the extracellular proteins, are used to transform cells for production of recombinant DNA molecules. More generally the DNA from other pathogens can be used in vaccines, e.g. against other bacteria, viruses, fungi and protozoa. Since different combinations of DNA can be used, a wide range of effective compositions can be produced. They generate a response against the antigens most often found on infected cells during the infection, regardless of the strength or

CC	specificity of the immune response. The vaccines are easy to produce
CC	and less toxic than known killed or attenuated vaccines, so can be given
CC	to immunocompromised subjects, e.g. those with HIV infection.
XX	
SQ	Sequence 325 AA;
	Query Match.
	Best local Similarity 100.0%; Score 1745; DB 18; Length 325;
	Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MTDSRKIRAWGRRLMIGTAAAVVLGVLGAGGAATAGAFSRPGLPVLYLVQVPSMGRR 60
Db	1 MTDSRKIRAWGRRLMIGTAAAVVLGVLGAGGAATAGAFSRPGLPVLYLVQVPSMGRR 60
QY	61 DIKVFQSGGNNSPAVYLIDGLRAQDDYNGWDINTPAFEWYYQSGLSIVMPVGQSSFFYS 120
Db	61 DIKVFQSGGNNSPAVYLIDGLRAQDDYNGWDINTPAFEWYYQSGLSIVMPVGQSSFFYS 120
QY	121 DWTSPCAGCKAGCCTYKWETFLTSELQWLISANRAVKPTGSAATGLSMAGSSAMILAAHYH 180
Db	121 DWTSPCAGCKAGCCTYKWETFLTSELQWLISANRAVKPTGSAATGLSMAGSSAMILAAHYH 180
QY	181 QQFIYAGSLIALLDPSCQGMGPSLIIGLAMDAGGYKAADWMWGSPSDPAWRNPDTCQIPKL 240
Db	181 QQFIYAGSLIALLDPSCQGMGPSLIIGLAMDAGGYKAADWMWGSPSDPAWRNPDTCQIPKL 240
QY	241 VANNTRLWVCYCGNTGTNELGGANI PAEFTLENVRSSNLKFQDAYNAAGGHNAVFNFPNG 300
Db	241 VANNTRLWVCYCGNTGTNELGGANI PAEFTLENVRSSNLKFQDAYNAAGGHNAVFNFPNG 300
QY	301 THSWFYWGQLNAMKGDLQSSLGAG 325
Db	301 THSWFYWGQLNAMKGDLQSSLGAG 325

RESULT 2	
AAW63032	
ID	AAW63032 standard; Protein; 325 AA.
XX	
AAW63032;	
XX	
DT	23-OCT-1998 (first entry)
XX	
XX	Mycobacterium tuberculosis 30 kD protein sequence.
XX	
KW	Mycobacterium tuberculosis; vaccination; extracellular product;
KW	immunodominant epitope; interleukin-12; MF59; immune response;
KW	opsonising humoral response; intracellular pathogen.
XX	
OS	Mycobacterium tuberculosis.
XX	
PN	WO9831388-A1.
XX	
PD	23-JUL-1998.
XX	
PF	15-JAN-1998; 98WO-US00942.
XX	
PR	21-JAN-1997; 97US-0786533.
XX	
PA	(REGC) UNIV CALIFORNIA.
XX	
PI	Harth G, Horwitz MA, Lee B;
XX	
DR	WPI; 1998-413815/35.
XX	N-PSDB; AAV42595.
XX	
PT	Vaccines against Mycobacterium containing major extracellular
PT	proteins - used to, e.g. induce protective and therapeutic immune
PT	responses, and for detecting an immune response
XX	
PS	Example 2; Pages 17-38; 236pp; English.
XX	
CC	This represents a Mycobacterium tuberculosis 30 kD protein. The invention

CC provides an agent for vaccinating mammals against Mycobacterium. The
 CC agent comprises at least one of the major abundant extracellular 110,
 CC 80, 71, 58, 45, 32A, 32B, 30, 24, 23.5, 23, 16, 14 or 12 KD proteins of
 CC M. tuberculosis, or at least 1 of their immunodominant epitopes and
 CC interleukin-12 (IL-12) or MF59 as adjuvants. The agent containing the
 CC nucleic acid encoding the extracellular products are used to raise a
 CC protective or therapeutic immune response against Mycobacterium,
 CC specifically M. tuberculosis. The immunodominant epitopes can also be
 CC used (typically in a cutaneous hypersensitivity test) to detect an
 CC immune response to vaccination. Preparation of the agent does not require
 CC selection of the most immunogenic products, so large scale production and
 CC purification are easy, resulting in a consistent, standardised
 CC formulation, having lower toxicity than killed or attenuated vaccines.
 CC The agents provide a rapid and effective response (including a strong
 CC cell-mediated component) and are safe even in immunocompromised subjects.
 CC They prevent development of an opsonising humoral response that might
 CC spread intracellular pathogens.

XX SQ Sequence 325 AA;
 Query Match 100.0%; Score 1745; DB 19; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.2e-143;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGATAGAFSRPGLPVEYLQVPSMGR 60
 DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGATAGAFSRPGLPVEYLQVPSMGR 60
 QY 61 DIKVPQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
 DB 61 DIKVPQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
 QY 121 DWSPACGKAGCCTYKWTFTLSELPQWLSANRAVKPTGSAATGLSMAGSSAMILAAHP 180
 DB 121 DWSPACGKAGCCTYKWTFTLSELPQWLSANRAVKPTGSAATGLSMAGSSAMILAAHP 180
 QY 181 QQFIYAGSLALLDPQSGMPSLIGLAWGDAGGYKAADMWGPSSDPAWERNDDTQIIPKL 240
 DB 181 QQFIYAGSLALLDPQSGMPSLIGLAWGDAGGYKAADMWGPSSDPAWERNDDTQIIPKL 240
 QY 241 VANNTLWYVCGNTPNELGGANI PAEFLNFVRSSNLKFDQDAYNAAGGHNAVFNPFPNG 300
 DB 241 VANNTLWYVCGNTPNELGGANI PAEFLNFVRSSNLKFDQDAYNAAGGHNAVFNPFPNG 300
 QY 301 THSWEYWGALNMGKDLQSSLGAG 325
 DB 301 THSWEYWGALNMGKDLQSSLGAG 325

RESULT 3

AAW46776
 ID AAW46776 standard; peptide; 325 AA.
 AC AAW46776;
 XX
 XX 09-JUL-1998 (first entry)
 DE Amino acid sequence of alpha antigen of Mycobacterium tuberculosis.
 XX Alpha antigen; mycobacterial disease; monoclonal antibody; detection;
 KW Mycobacteria; Mycobacterium avium; intracellular.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO9804918-A1.
 XX
 XX 05-FEB-1998.
 XX
 PF 24-JUN-1997; 97WO-US13235.
 XX
 PR 26-JUL-1996; 96US-0690347.
 XX
 PA (UYCA-) UNIV CASE WESTERN RESERVE.

XX Wallis RS;
 DR WPI; 1998-130832/12.
 XX
 PT Detection of mycobacteria - using a monoclonal antibody directed
 PT against a portion of the alpha antigen, and detection of complexes
 PT formed
 XX
 PS Claim 11; Fig 19; 105pp; English.

XX The present sequence represents the alpha antigen of Mycobacterium
 CC tuberculosis. The alpha antigen is an abundantly produced, secreted 30
 CC kilodalton mycobacterial protein involved in mycolic acid synthesis. It
 CC is an useful indicator of mycobacterial disease as it can account for up
 CC to 20% of the protein content of spent mycobacterial culture medium. In
 CC addition, homologous alpha proteins exist for most species of
 CC Mycobacteria. Monoclonal antibodies were produced against the present
 CC sequence and its fragments. These can be used for the detection of
 CC Mycobacteria in a test sample. The method can be used for the detection
 CC of Mycobacteria such as M. tuberculosis, M. avium and M. intracellulare.
 CC The method can be used to monitor the progression of therapy in
 CC individuals infected with Mycobacteria.

XX SQ Sequence 325 AA;

Query Match 100.0%; Score 1745; DB 19; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.2e-143;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGATAGAFSRPGLPVEYLQVPSMGR 60
 DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGATAGAFSRPGLPVEYLQVPSMGR 60
 QY 61 DIKVPQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
 DB 61 DIKVPQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
 QY 121 DWSPACGKAGCCTYKWTFTLSELPQWLSANRAVKPTGSAATGLSMAGSSAMILAAHP 180
 DB 121 DWSPACGKAGCCTYKWTFTLSELPQWLSANRAVKPTGSAATGLSMAGSSAMILAAHP 180
 QY 181 QQFIYAGSLALLDPQSGMPSLIGLAWGDAGGYKAADMWGPSSDPAWERNDDTQIIPKL 240
 DB 181 QQFIYAGSLALLDPQSGMPSLIGLAWGDAGGYKAADMWGPSSDPAWERNDDTQIIPKL 240
 QY 241 VANNTLWYVCGNTPNELGGANI PAEFLNFVRSSNLKFDQDAYNAAGGHNAVFNPFPNG 300
 DB 241 VANNTLWYVCGNTPNELGGANI PAEFLNFVRSSNLKFDQDAYNAAGGHNAVFNPFPNG 300
 QY 301 THSWEYWGALNMGKDLQSSLGAG 325
 DB 301 THSWEYWGALNMGKDLQSSLGAG 325

RESULT 4

AAV14853
 ID AAV14853 standard; protein; 325 AA.
 XX
 AC AAV14853;
 XX
 XX 25-OCT-1999 (first entry)
 DE Antigen 85B protein from M. tuberculosis.
 XX
 XX Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
 KW dendritic cell maturation; infectious disease; immune disorder; cancer;
 KW respiratory system; mycobacterial infection; allergy; tuberculosis;
 KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
 KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
 XX squamous cell carcinoma; melanoma.
 XX
 OS Mycobacterium tuberculosis.


```

Db 121 DWSPACGKAGCQTYKWEFTLTSELPOWLSANRAVKPTGSAAGLSWAGSSAMTILAAAYHP 180
QY 181 QQFIYAGSLSALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNDDPTQQIPKL 240
Db 181 QQFIYAGSLSALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNDDPTQQIPKL 240
QY 241 VANNTRLWVYCGNGTNPNEGANI PAEFLFNFRSSNLKFDQDAYNAAGHNAVFPPNG 300
Db 241 VANNTRLWVYCGNGTNPNEGANI PAEFLFNFRSSNLKFDQDAYNAAGHNAVFPPNG 300
QY 301 THSWEYWGAGQLNAMKGDQLSSLGAG 325
Db 301 THSWEYWGAGQLNAMKGDQLSSLGAG 325

RESULT 6
ABB73459
ID ABB73459 standard; Protein; 325 AA.
XX
AC ABB73459;
XX
DT 08-APR-2002 (first entry)
XX
DE M tuberculosis 85B protein SEQ ID NO: 33.
XX
KW Skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic;
KW Th2 immune response; immunomodulatory.
XX
OS Mycobacterium tuberculosis.
XX
PN US6328978-B1.
XX
PD 11-DEC-2001.
XX
PF 02-JUN-1999; 99US-0324542.
XX
PR 23-DEC-1997; 97US-0997080.
XX
PA (GENE-) GENESIS RES & DEV CORP LTD.
XX
PI Watson JD, Tan PLJ, Prestidge R;
XX
WP1; 2002-138361/18.
XX
Inhibiting skin inflammation associated with skin disorder e.g.
PT psoriasis, by administering composition comprising delipidated and
PT deglycolipidated Mycobacterium vaccae cells or Mycobacterium vaccae
PT culture filtrate -
XX
Example 5; Column 61-62; 116pp; English.
XX
The present invention relates to a method of inhibiting skin inflammation
CC associated with a skin disorder selected from psoriasis, atopic
CC dermatitis and allergic contact dermatitis, which involves administering
CC a composition containing delipidated and deglycolipidated Mycobacterium
CC vaccae cells or M. vaccae culture filtrate. The skin disorder to be
CC treated may also include alopecia areata, and skin cancers such as basal
CC cell carcinoma, squamous cell carcinoma and melanoma. The composition
CC acts by inhibiting the Th2 immune response. The present sequence is a
CC protein described in the exemplification of the invention.
XX
SQ Sequence 325 AA;
Query Match 100.0%; Score 1745; DB 23; Length 325;
Best Local Similarity 100.0%; Pred. No. 1.2e-143;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTDVSRKIRAWGRRLMIGTAAAVVLGVLGAGTAAGAFGRPLGVYLOVPSMGR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVVLGVLGAGTAAGAFGRPLGVYLOVPSMGR 60

```

```

QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGQSSFFS 120
Db 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGQSSFFS 120
QY 121 DWSPACGKAGCQTYKWEFTLTSELPOWLSANRAVKPTGSAAGLSWAGSSAMTILAAAYHP 180
Db 121 DWSPACGKAGCQTYKWEFTLTSELPOWLSANRAVKPTGSAAGLSWAGSSAMTILAAAYHP 180
QY 181 QQFIYAGSLSALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNDDPTQQIPKL 240
Db 181 QQFIYAGSLSALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNDDPTQQIPKL 240
QY 241 VANNTRLWVYCGNGTNPNEGANI PAEFLFNFRSSNLKFDQDAYNAAGHNAVFPPNG 300
Db 241 VANNTRLWVYCGNGTNPNEGANI PAEFLFNFRSSNLKFDQDAYNAAGHNAVFPPNG 300
QY 301 THSWEYWGAGQLNAMKGDQLSSLGAG 325
Db 301 THSWEYWGAGQLNAMKGDQLSSLGAG 325

RESULT 7
AAM50730
ID AAM50730 standard; Protein; 325 AA.
XX
AC AAM50730;
XX
DT 18-APR-2002 (first entry)
XX
DE Mycobacterium tuberculosis immunodominant Mtb protein FbpB Ag85B.
XX
KW Mtb; FbpB; Ag85B; immunogen; mycobacteria; immunisation; vaccine.
XX
OS Mycobacterium tuberculosis.
XX
FH Key Location/Qualifiers
FT Peptide 1..40
FT /label= Signal peptide
FT /note= "SEC-dependent signal secretion sequence"
FT Protein 41..325
FT /label= Mature protein
XX
PN WO200204018-A2.
XX
PD 17-JAN-2002.
XX
PF 10-JUL-2001; 2001WO-US21717.
XX
PR 10-JUL-2000; 2000US-217646P.
XX
PA (COLS ) UNIV COLORADO STATE RES FOUND.
XX
PI Orme IM, Belisle JT;
XX
WP1; 2002-164602/21.
XX
Vaccine for boosting immunity to mycobacteria when administered in
PT mid-life in a subject who has been vaccinated in childhood with
PT Bacillus Calmette-Guerrin, has purified proteins from mycobacterium
PT tuberculosis -
XX
Claim 8; Page 17; 61pp; English.
XX
The present sequence is that of the Mycobacterium tuberculosis
CC (Mtb) strain H37Rv gene Rv1886c product, designated FbpB Ag85B.
CC This is one of 31 immunodominant secreted or cytosolic Mtb proteins
CC of strain H37Rv (see AAM50729-59) discovered through the use of
CC 2-dimensional liquid phase electrophoresis coupled with an in vitro
CC interferon-gamma assay and liquid chromatography-mass spectrometry.
CC The immunogens stimulate a strong interferon-gamma response from
CC T cells of M. tuberculosis infected mice. The invention provides
CC vaccine compositions for boosting immunity to mycobacteria when

```

CC administered in mid-life to a subject who has been vaccinated
 CC neonatally or in early childhood with BCG and in whom protective
 CC immunity has waned. The vaccine compositions comprise 1 or more
 CC of the 31 purified immunogenic proteins. When used as immunogens,
 CC the secreted Mtb proteins lack the secreted signal sequence. A
 CC preferred protein is Ag85A (see AAM50759), the secreted product of
 CC the RV3084v gene.

XX SQ Sequence 325 AA;

Query Match 100.0%; Score 1745; DB 23; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.2e-143;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFRRPGLPVEYLQVPSMGR 60
 Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFRRPGLPVEYLQVPSMGR 60
 Qy 61 DIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFS 120
 Db 61 DIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFS 120
 Qy 121 DWYSPACGKAGCQYKWTETLTSELPOWLSANRAVKPTGSAIIGLSMAGSSAMILAAVHP 180
 Db 121 DWYSPACGKAGCQYKWTETLTSELPOWLSANRAVKPTGSAIIGLSMAGSSAMILAAVHP 180
 Qy 181 QQFTYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPPOWERNDPTQOIPKL 240
 Db 181 QQFTYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPPOWERNDPTQOIPKL 240
 Qy 241 VANNTRLWVYCGNGTPELNGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVFPPNG 300
 Db 241 VANNTRLWVYCGNGTPELNGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVFPPNG 300
 Qy 301 THSWEYGAQLNAMKGDLOSSLGAG 325
 Db 301 THSWEYGAQLNAMKGDLOSSLGAG 325

RESULT 8
 ABUS6406
 ID ABUS6406 standard; Protein; 325 AA.
 AC ABUS6406;
 XX
 DT 31-MAR-2003 (first entry)
 XX
 DE Mycobacterium tuberculosis 30kDa protein.
 KW Antibacterial; tuberculosis; vaccine; gene therapy; 30kDa protein.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN US2002131975-A1.
 XX
 PD 19-SEP-2002.
 XX
 PF 14-SEP-2001; 2001US-0953510.
 XX
 PR 23-MAY-1996; 96US-0652842.
 PR 23-NOV-1993; 93US-0156358.
 PR 12-AUG-1994; 94US-0289667.
 PR 23-MAY-1995; 95US-0447398.
 PR 31-OCT-1995; 95US-0551149.
 PR 06-DEC-1996; 96US-0568357.
 PR 21-SEP-1998; 98US-0157689.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Horwitz MA, Earth G;
 XX
 DR WPI; 2003-174073/17.
 DR N-PSDB; ABX75988.

XX New vaccine, useful for promoting an immune response against infectious
 PT pathogens of the genus Mycobacterium in a mammalian host -
 XX
 PS Example 1; Page 14-15; 82pp; English.

XX The invention describes a vaccine for promoting an immune response, in
 CC a mammalian host, against infectious pathogens of the genus
 CC Mycobacterium, comprising at least 1 immunodominant epitope of at least
 CC one majority abundant extracellular product comprising Mycobacterium
 CC tuberculosis 110, 80, 71, 58, 45, 32A, 32B, 30, 24, 23.5, 23, 16, 14,
 CC 12 KD proteins or their analogues, homologues or subunits.
 CC The proteins and polypeptides of the invention are useful in gene
 CC therapy and treatment of diseases caused by Mycobacterium such as
 CC tuberculosis. This is the amino acid sequence of the Mycobacterium
 CC tuberculosis 30kDa protein.

XX SQ Sequence 325 AA;

Query Match 100.0%; Score 1745; DB 24; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.2e-143;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFRRPGLPVEYLQVPSMGR 60
 Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFRRPGLPVEYLQVPSMGR 60
 Qy 61 DIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFS 120
 Db 61 DIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFS 120
 Qy 121 DWYSPACGKAGCQYKWTETLTSELPOWLSANRAVKPTGSAIIGLSMAGSSAMILAAVHP 180
 Db 121 DWYSPACGKAGCQYKWTETLTSELPOWLSANRAVKPTGSAIIGLSMAGSSAMILAAVHP 180
 Qy 181 QQFTYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPPOWERNDPTQOIPKL 240
 Db 181 QQFTYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPPOWERNDPTQOIPKL 240
 Qy 241 VANNTRLWVYCGNGTPELNGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVFPPNG 300
 Db 241 VANNTRLWVYCGNGTPELNGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVFPPNG 300
 Qy 301 THSWEYGAQLNAMKGDLOSSLGAG 325
 Db 301 THSWEYGAQLNAMKGDLOSSLGAG 325

RESULT 9
 ABB73461
 ID ABB73461 standard; Protein; 323 AA.
 XX
 AC ABB73461;
 XX
 DT 08-APR-2002 (first entry)
 XX
 DE M bovis 85B protein SEQ ID NO: 35.
 XX
 KW Skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
 KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
 KW antipsoriatic; dermatological; antiinflammatory; antiallergic;
 XX Th2 immune response; immunomodulatory.
 OS Mycobacterium bovis.
 XX
 PN US6328978-B1.
 XX
 PD 11-DEC-2001.
 XX
 PF 02-JUN-1999; 99US-0324542.
 XX
 PR 23-DEC-1997; 97US-0997080.
 XX

(GENE-) GENESIS RES & DEV CORP LTD.
 Watson JD, Tan PLJ, Prestidge R;
 WPI; 2002-138361/18.
 Inhibiting skin inflammation associated with skin disorder e.g. psoriasis, by administering composition comprising delipidated and deglycolipidated Mycobacterium vaccae cells or Mycobacterium vaccae culture filtrate
 Example 5; Column 65-66; 116pp; English.
 The present invention relates to a method of inhibiting skin inflammation associated with a skin disorder selected from psoriasis, atopic dermatitis and allergic contact dermatitis, which involves administering a composition containing delipidated and deglycolipidated Mycobacterium vaccae cells or M. vaccae culture filtrate. The skin disorder to be treated may also include alopecia areata, and skin cancers such as basal cell carcinoma, squamous cell carcinoma and melanoma. The composition acts by inhibiting the Th2 immune response. The present invention is a protein described in the exemplification of the invention.
 SQ Sequence 323 AA;
 Query Match 97.8%; Score 1706; DB 23; Length 323;
 Best Local Similarity 98.5%; Pred. No. 3e-140;
 Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
 QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
 DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
 QY 61 DIKVFQSGGNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
 DB 61 DIKVFQSGGNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
 QY 121 DWSPACGKAGCQTYKWTFTLSELPAWLSANRAVKPTGSAALGLSWAGSSAMILAAHP 180
 DB 121 DWSPACGKAGCQTYKWTFTLSELPAWLSANRAVKPTGSAALGLSWAGSSAMILAAHP 180
 QY 181 QQFIYAGSLALLDPSQGMG--LIGLAMGAGGYKAADMWGPSSDPWERNNDPTQIIPKL 240
 DB 181 QQFIYAGSLALLDPSQGMG--LIGLAMGAGGYKAADMWGPSSDPWERNNDPTQIIPKL 238
 QY 241 VANNRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDAYNAAGGHNAVFNPFG 300
 DB 239 VANNRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDAYNAAGGHNAVFNPFG 298
 QY 301 THSWEYWGAGLNAMKGDLSLQSLGAG 325
 DB 299 THSWEYWGAGLNAMKGDLSLQSLGAG 323
 RESULT 10
 AAY14855
 ID AAY14855 standard; protein; 323 AA.
 AC AAY14855;
 DT 25-OCT-1999 (first entry)
 XX Antigen 85B protein from M. bovis.
 XX Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
 KW dendritic cell maturation; infectious disease; immune disorder; cancer;
 KW respiratory system; mycobacterial infection; allergy; tuberculosis;
 KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
 KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
 XX squamous cell carcinoma; melanoma.
 OS Mycobacterium bovis.
 XX

PN WO9932634-A2.
 XX 01-JUL-1999.
 XX 23-DEC-1998; 98WO-NZ00189.
 XX 04-DEC-1998; 98US-0205426.
 PR 23-DEC-1997; 97US-0996624.
 PR 23-DEC-1997; 97US-0997080.
 PR 23-DEC-1997; 97US-0997362.
 PR 11-JUN-1998; 98US-0095855.
 PR 17-SEP-1998; 98US-0156181.
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 XX Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
 WPI; 1999-430163/36.
 XX Enhancing immune response to an antigen
 XX Example 11; Page 158-159; 243pp; English.
 CC The invention provides heat-killed Mycobacterium vaccae, or recombinant
 CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
 CC T cells and natural killer cells, to stimulate the production of
 CC cytokines, to enhance the expression of co-stimulatory molecules on
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation
 CC and function. The proteins can be expressed by standard recombinant
 CC methodology. Pharmaceutical compositions comprising the proteins or
 CC nucleic acid sequences encoding the proteins can be used for the
 CC treatment, prevention, and detection of disorders including infectious
 CC diseases, immune disorders and cancer. In particular, the compounds and
 CC methods are used for treatment of diseases of the respiratory system,
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
 CC carcinoma and melanoma.
 XX SQ Sequence 323 AA;
 Query Match 97.4%; Score 1699; DB 20; Length 323;
 Best Local Similarity 98.2%; Pred. No. 1.2e-139;
 Matches 319; Conservative 0; Mismatches 4; Indels 2; Gaps 1;
 QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
 DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
 QY 61 DIKVFQSGGNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
 DB 61 DIKVFQSGGNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
 QY 121 DWSPACGKAGCQTYKWTFTLSELPAWLSANRAVKPTGSAALGLSWAGSSAMILAAHP 180
 DB 121 DWSPACGKAGCQTYKWTFTLSELPAWLSANRAVKPTGSAALGLSWAGSSAMILAAHP 180
 QY 181 QQFIYAGSLALLDPSQGMG--LIGLAMGAGGYKAADMWGPSSDPWERNNDPTQIIPKL 240
 DB 181 QQFIYAGSLALLDPSQGMG--LIGLAMGAGGYKAADMWGPSSDPWERNNDPTQIIPKL 238
 QY 241 VANNRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDAYNAAGGHNAVFNPFG 300
 DB 239 VANNRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDAYNAAGGHNAVFNPFG 298
 QY 301 THSWEYWGAGLNAMKGDLSLQSLGAG 325
 DB 299 THSWEYWGAGLNAMKGDLSLQSLGAG 323
 RESULT 11
 AAR10486

```

ID  AAR10486 standard; Protein; 325 AA.
XX  AAR10486;
AC  15-APR-1991 (first entry)
XX  Tuberculin active protein alpha-antigen.
XX  pneumonia; lung cancer; BCG live vaccine; tuberculin.
XX  Mycobacterium kansasii.
XX  Key Location/Qualifiers
XX  Peptide 1..40
XX  /label= signal peptide
XX  Protein 41..325
XX  /label= Tuberculin active protein alpha-antigen
XX  JP02308793-A.
XX  21-DEC-1990.
XX  22-MAY-1989; 89JP-0128091.
XX  22-MAY-1989; 89JP-0128091.
XX  (AJIN ) AJINOMOTO KK.
XX  WPI; 1991-041063/06.
XX  N-PSDB; AAQ10372.
XX  Alpha antigen derived from mycobacterium kansasii - for diagnosis
XX  of mycobacterium kansasii from pneumonia and lung cancer and for
XX  development of BCG vaccine
XX  Claim 1; Page 1; 11pp; Japanese.
XX  Alpha-antigen is produced in large quantities by culturing, under
XX  aerobic conditions, host cells that have been transformed with
XX  recombinant DNA comprising the sequence encoding the mature alpha-
XX  antigen and its signal sequence. The protein distinguishes
XX  M.kansasii disease from pneumonia or lung cancer.
XX  SQ Sequence 325 AA;
Query Match 91.8%; Score 1602; DB 12; Length 325;
Best Local Similarity 89.5%; Pred. No. 3.4e-131;
Matches 290; Conservative 20; Mismatches 14; Indels 0; Gaps 0;
QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSGKIRAWGRRLLVGAAAAALPGLVGLAGGAATAGAFSRPGLPVEYHQPVAAMGR 60
QY 61 DIKVOFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
DB 61 SKVQFQSGGDNPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
QY 121 DWYPACGKAGCQTYKWTFTLSELPMLSANRVPKPTGSAAGTSMAGSSAMILAAVHP 180
DB 121 DWYPACGKAGCQTYKWTFTLSELPMLSANRVPKPTGSAAGTSMAGSSAMILAAVHP 180
QY 181 QOPIYAGLSALLDPSQGMPSLIGLAMDAGGYKAADMWGPSSDPAMERNPTQOIPKL 240
DB 181 QOPIYAGLSALLDPSQGMPSLIGLAMDAGGYKAADMWGPSSDPAMERNPTQOIPKL 240
QY 241 VANNTRLWYCCNGTPELGGANIIPAEFLNFRSSNLKFDQAYNAAGHNAVFNPFG 300
DB 241 VANNTRLWYCCNGTPELGGANIIPAEFLNFRSSNLKFDQAYNAAGHNAVFNPFG 300
QY 301 THSWEYWGACLNAKMGDLOSSIGA 324
DB 301 THSWEYWGACLNAKMGDLOSSIGA 324

```

```

RESULT 12
AAR08099
ID  AAR08099 standard; protein; 325 AA.
XX  AAR08099;
AC  01-MAR-1991 (first entry)
XX  Myconacterium derived alpha-antigen gene product.
XX  Vaccine; tuberculosis; HIV-1.
XX  Mycobacterium kansasii.
XX  Key Location/Qualifiers
XX  Peptide 1..40
XX  /label= Signal peptide
XX  Protein 41..325
XX  /label= Mature protein
XX  EP400973-A.
XX  05-DEC-1990.
XX  30-MAY-1990; 90EP-0305849.
XX  16-MAR-1990; 90JP-0064310.
XX  31-MAY-1989; 89JP-0135855.
XX  (AJIN ) AJINOMOTO KK.
XX  Matsuo K, Yamaguchi R, Yamazaki A, Yamada T;
XX  WPI; 1990-363461/49.
XX  N-PSDB; AAQ06668.
XX  New mycobacterial secretory vector - used to transform host
XX  cells, and vaccine comprising transformant
XX  Disclosure; Fig 1; 19pp; English.
XX  The signal peptide may be used to produce a recombinant protein
XX  comprising alpha-antigen and B-cell epitope of the HIV-1 gag
XX  antigen p17.
XX  The fused product is useful in vaccine production eg. tuberculosis.
XX  A live vaccine such as M.bovis BCG or M.smegmatis has a low toxicity
XX  and long lasting activity.
XX  SQ Sequence 325 AA;
Query Match 91.4%; Score 1595; DB 11; Length 325;
Best Local Similarity 89.2%; Pred. No. 1.4e-130;
Matches 289; Conservative 20; Mismatches 15; Indels 0; Gaps 0;
QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSGKIRAWGRRLLVGAAAAALPGLVGLAGGAATAGAFSRPGLPVEYLQVPSAAMGR 60
QY 61 DIKVOFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
DB 61 SKVQFQSGGDNPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
QY 121 DWYPACGKAGCQTYKWTFTLSELPMLSANRVPKPTGSAAGTSMAGSSAMILAAVHP 180
DB 121 DWYPACGKAGCQTYKWTFTLSELPMLSANRVPKPTGSAAGTSMAGSSAMILAAVHP 180
QY 181 QOPIYAGLSALLDPSQGMPSLIGLAMDAGGYKAADMWGPSSDPAMERNPTQOIPKL 240
DB 181 QOPIYAGLSALLDPSQGMPSLIGLAMDAGGYKAADMWGPSSDPAMERNPTQOIPKL 240
QY 241 VANNTRLWYCCNGTPELGGANIIPAEFLNFRSSNLKFDQAYNAAGHNAVFNPFG 300
DB 241 VANNTRLWYCCNGTPELGGANIIPAEFLNFRSSNLKFDQAYNAAGHNAVFNPFG 300

```

Db	241	VANNRLTWYCGNGTSPSELGGANVPAPFLNFVRSSNLKFDQDAYNAAGGHNAVFNLDANG	300
QY	301	THSWEYWGQALNAMKGDLOSSLGA	324
Db	301	THSWEYWGQALNAMKGDLOSSLGA	324
RESULT 13			
AAW72943	ID	AAW72943 standard; Protein; 403 AA.	
XX	AC	AAW72943;	
XX	XX	21-JAN-1999 (first entry)	
DE	XX	Mycobacterium tuberculosis antigen ESAT6-MPT59.	
DE	XX	Mycobacterium tuberculosis; antigen; vaccine; immunological;	
KW	KW	immunogen; infection.	
XX	XX	Mycobacterium tuberculosis.	
OS	XX	WO9844119-A1.	
PN	XX	08-OCT-1998.	
PD	XX	01-APR-1998; 98WO-DK00132.	
PF	XX	05-JAN-1998; 98US-0070488.	
PR	PR	12-APR-1997; 97DK-0000376.	
PR	PR	18-APR-1997; 97US-0044624.	
PR	PR	10-NOV-1997; 97DK-0001277.	
XX	XX	(STAT-) STATENS SERUM INST.	
PA	PI	Andersen P, Florio W, Nielsen R, Oettinger T, Rasmussen PB;	
PI	PI	Rosenkrands I, Weldingh K;	
XX	XX	WPI; 1998-542705/46.	
DR	XX	New isolated mycobacteria polypeptides and nucleic acids - used for	
PT	PT	developing products for the diagnosis of or vaccination against	
PT	PT	mycobacterial infections, particularly tuberculosis	
XX	XX	Disclosure; Page 233-234; 163pp; English.	
ES	XX	The present sequence represents a Mycobacterium tuberculosis protein.	
CC	CC	Products from the present invention, which describes protein fragments	
CC	CC	and nucleic acid fragments derived from M.tuberculosis, can be used in	
CC	CC	the detection of and prevention of mycobacterial infections. In	
CC	CC	particular, the proteins and nucleic acids can be used for the diagnosis	
CC	CC	of or vaccination against tuberculosis caused by M. tuberculosis,	
CC	CC	M. africanum or M. bovis.	
XX	XX	Sequence 403 AA;	
QY	Query Match	88.8%; Score 1550; DB 19; Length 403;	
Db	Best Local Similarity	99.3%; Pred. No. 1.5e-126;	
QY	Matches 286; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
QY	38	AGAFSRPGLPVEYLQVPSMGRDIKVQFOSGGNNSPAVLLDGLRAQDDYNGWDINTPA	97
Db	116	AKLFSRREL PVEYLQVPSMGRDIKVQFOSGGNNSPAVLLDGLRAQDDYNGWDINTPA	175
QY	98	FEWYQSGLSIVMPVGQSSFYSDWYSPACGKAGCQTYKWTFTLTSELPOWLSANRAVKP	157
Db	176	FEWYQSGLSIVMPVGQSSFYSDWYSPACGKAGCQTYKWTFTLTSELPOWLSANRAVKP	235
QY	158	TGSAATGLSNAGSSAMILAAHYHQQFYIAGLSALLDPSQGMGPSLTGLAMGAGGYKAA	217
Db	236	TGSAATGLSNAGSSAMILAAHYHQQFYIAGLSALLDPSQGMGPSLTGLAMGAGGYKAA	295
QY	218	DMWGSPSDPAWERNDDPTQIQPKLVANNTRLWTVCGNGT PNELGANI PAEFLNFVRSSN	277

DR N-PSDB; AAN90130.
XX Alpha-antigen derived from BCG bacteria - used for diagnosing
PT tuberculous cephalomeningitis.
XX
XX
PS Claim 1; page 576; pp; Japanese.
XX
XX Tuberculin active protein alpha-antigen is derived from Mycobacterium
CC bovis (see AAN90130). The protein is useful as a pharmaceutical agent
CC for diagnosing tuberculous cephalomeningitis.
XX
SQ Sequence 283 AA;

Query Match 86.5%; Score 1510; DB 10; Length 283;
Best Local Similarity 98.2%; Pred. No. 3e-123;
Matches 280; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 41 FSRGLPVEYLQVPSMGRDIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEW 100
Db 1 FSRGLPVEYLQVPSMGRDIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEW 60
QY 101 YYQSGLSIVMPVGGQSPSYDWYSPACGACQTYKWTFTLSELPOWLSANRAVKPTGS 160
Db 61 YYQSGLSIVMPVGGQSPSYDWYSPACGACQTYKWTFTLSELPOWLSANRAVKPTGS 120
QY 161 AAIGLSMAGSSAMILAAVHPQOPIYAGSLSALLDPSQGMGLIGLAMGDAGGYKAADWW 220
Db 121 AAIGLSMAGSSAMILAAVHPQOPIYAGSLSALLDPSQGMGLIGLAMGDAGGYKAADWW 178
QY 221 GPSSDPANERNDPTQOIPKLVANNTRLWVYCGNGTPNELGGANIPAEFLENFVRSSNLKF 280
Db 179 GPSSDPANERNDPTQOIPKLVANNTRLWVYCGNGTPNELGGANIPAEFLENFVRSSNLKF 238
QY 281 QDAYNAGGHNAVFNFPNGTHSWEYWGQAQLNAMKGDILQSSILGAG 325
Db 239 QDAYNAGGHNAVFNFPNGTHSWEYWGQAQLNAMKGDILQSSILGAG 283

RESULT 19
AAY14851
ID AAY14851 standard; protein; 327 AA.
XX
AC AAY14851;
XX
DT 25-OCT-1999 (first entry)
XX
DE Antigen 85B protein from M. leprae.
XX
KW Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
KW dendritic cell maturation; infectious disease; immune disorder; cancer;
KW respiratory system; mycobacterial infection; allergy; tuberculosis;
KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
KW squamous cell carcinoma; melanoma.
XX
OS Mycobacterium leprae.
XX
PN W09932634-A2.
XX
PD 01-JUL-1999.
XX
PF 23-DEC-1998; 98WO-NZ00189.
XX
PR 04-DEC-1998; 98US-0205426.
PR 23-DEC-1997; 97US-0996624.
PR 23-DEC-1997; 97US-0997080.
PR 23-DEC-1997; 97US-0997362.
PR 11-JUN-1998; 98US-0095855.
PR 17-SEP-1998; 98US-0156181.
XX
XX (GENE-) GENESIS RES & DEV CORP LTD.
PA Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
XX
PI

XX WPI; 1999-430163/36.
DR
XX Enhancing immune response to an antigen
PT
XX
XX PS Example 11; Page 154-155; 243pp; English.
XX
XX The invention provides heat-killed Mycobacterium vaccae, or recombinant
CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
CC T cells and natural killer cells, to stimulate the production of
CC cytokines, to enhance the expression of co-stimulatory molecules on
CC dendritic cells and monocytes, and to enhance dendritic cell maturation
CC and function. The proteins can be expressed by standard recombinant
CC methodology. Pharmaceutical compositions comprising the proteins or
CC nucleic acid sequences encoding the proteins can be used for the
CC treatment, prevention, and detection of disorders including infectious
CC diseases, immune disorders and cancer. In particular, the compounds and
CC methods are used for treatment of diseases of the respiratory system,
CC such as mycobacterial infections, asthma, allergies, tuberculosis,
CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
CC carcinoma and melanoma.
SQ Sequence 327 AA;

Query Match 83.3%; Score 1454; DB 20; Length 327;
Best Local Similarity 83.0%; Pred. No. 2.7e-118;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLVAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MIDVSGKIRAWGRLLVG--AAATLPSLISLAGGAATAAFSRPGLPVEYLQVPSMGR 58
QY 61 DIKVOFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSPSY 120
Db 59 TIKVQFQNGGSPAVYLLDGLRAQDDYNGWDINTSAFEWYQSGLSIVMPVGGQSPSY 118
QY 121 DWYSPACGACQTYKWTFTLSELPOWLSANRAVKPTGSAAIGLSMAGSSAMILAAVHP 180
Db 119 DWYSPACGACQTYKWTFTLSELPOWLSANRAVKPTGSAAIGLSMAGSSAMILAAVHP 178
QY 181 QQFIYAGSLSALLDPSQGMGLIGLAMGDAGGYKAADWWGPPSSPAWERNDPTQOIPKL 240
Db 179 DQFIYAGSLSALLDPSQGMGLIGLAMGDAGGYKAADWWGPPNDPAWORNDPILQAGKL 238
QY 241 VANNTRLWVYCGNGTPNELGGANIPAEFLENFVRSSNLKFQDAYNAGGHNAVFNFPNG 300
Db 239 VANNTRLWVYCGNGTPNELGGANIPAEFLENFVRSSNLKFQDAYNAGGHNAVFNFPNG 298
QY 301 THSWEYWGQAQLNAMKGDILQSSILGA 324
Db 299 THSWEYWGQAQLNAMKGDILQNTLMA 322

RESULT 20
ABB73457
ID ABB73457 standard; Protein; 327 AA.
XX
AC ABB73457;
XX
DT 08-APR-2002 (first entry)
XX
DE M leprae 85B protein SEQ ID NO: 31.
XX
KW Skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic;
KW Th2 immune response; immunomodulatory.
XX
OS Mycobacterium leprae.
XX
PN US6328978-B1.

XX 11-DEC-2001.
 PD 02-JUN-1999; 99US-0324542.
 PF 23-DEC-1997; 97US-0997080.
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA Watson JD, Tan PLJ, Prestidge R;
 PI WPI; 2002-138361/18.
 XX Inhibiting skin inflammation associated with skin disorder e.g.
 PT psoriasis, by administering composition comprising delipidated and
 PT deglycolipidated Mycobacterium vaccae cells or Mycobacterium vaccae
 PT culture filtrate
 XX Example 5; Column 57-60; 116pp; English.
 PS The present invention relates to a method of inhibiting skin inflammation
 CC associated with a skin disorder selected from psoriasis, atopic
 CC dermatitis and allergic contact dermatitis, which involves administering
 CC a composition containing delipidated and deglycolipidated Mycobacterium
 CC vaccae cells or M. vaccae culture filtrate. The skin disorder to be
 CC treated may also include alopecia areata, and skin cancers such as basal
 CC cell carcinoma, squamous cell carcinoma and melanoma. The composition
 CC acts by inhibiting the Th2 immune response. The present sequence is a
 CC protein described in the exemplification of the invention.
 XX Sequence 327 AA;
 SQ Query Match 83.3%; Score 1454; DB 23; Length 327;
 Best Local Similarity 83.0%; Pred. No. 2.7e-118;
 Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;
 QY 1 MTDVSKIRAWGRMLTGTAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
 DB 1 MIDVSGKIRAWGRWLLVG--AAATLPSLSLAGGAATASAFSRPGLPVEYLQVPSMGR 58
 QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFWYQSGLSIVMPVGGQSFYS 120
 DB 59 TIKVFQNGGNSPAVYLLDGLRAQDDYNGWDINTSAFWYQSGLSIVMPVGGQSFYS 118
 QY 121 DWTSPPAGKAGCCTYKWTETLTSELPOWLSANRAVKPTGSAATGLSMAGSSAMILAAHP 180
 DB 119 DWTSPPAGKAGCCTYKWTETLTSELPOWLSANRAVKPTGSAATGLSMAGSSAMILAAHP 178
 QY 181 QQFIYAGSLALDPQSGMGPSLIGLWAGDAGGYKAADMWGPSSDPAWERNDDTQIIPKL 240
 DB 179 DQFIYAGSLALMDSSQGIIPQLIGLWAGDAGGYKAADMWGPSSDPAWERNDDTQIIPKL 238
 QY 241 VANNTLWYVCGNGTPELNGGNIIPAEFLFNFRSSNLKFDQAYNAGGHNVAFFNPNG 300
 DB 239 VANNTLWYVCGNGTPELNGGNIIPAEFLFNFRSSNLKFDQAYNAGGHNVAFFNPNG 298
 QY 301 THSWEYWGALNMGDLOSSLGA 324
 DB 299 THSWEYWGALNMGDLOSSLGA 322
 RESULT 21
 ID ABB83662 standard; Protein; 285 AA.
 XX ABB83662;
 AC ABB83662;
 XX 15-NOV-2002 (first entry)
 DT Mycobacterium-originated alpha antigen.
 DE Antiallergic; Dermatological; Antiasthmatic; Antiinflammatory;
 XX Ophthalmological; Mycobacterium-originated alpha antigen;
 KW

KW allergic disease; atopic dermatitis; asthma; allergic rhinitis;
 KW allergic conjunctivitis; interleukin-4; Th2-type cytokine;
 XX IGE production; histamine; eosinophilic infiltration.
 OS Mycobacterium kansasii.
 XX WO200266055-A1.
 FN 29-AUG-2002.
 PD 20-FEB-2002; 2002WO-JP01459.
 PF 20-FEB-2001; 2001JP-0043291.
 XX (PRIM-) PRIMMUNE CORP INC.
 PA Yasutomi Y, Mizutani H;
 PI WPI; 2002-667038/71.
 XX Medicinal use of Mycobacterium-originated alpha-antigen or its gene in
 CC treating allergic diseases e.g. atopic dermatitis, asthma, allergic
 CC rhinitis and allergic conjunctivitis
 CC Disclosure; Page 27-28; 34pp; Japanese.
 XX This invention relates to pharmaceutical compositions for preventing
 CC or treating allergic diseases containing Mycobacterium-originated
 CC alpha antigens. The compositions are antiasthmatic, dermatological,
 CC antiasthmatic, antiinflammatory and ophthalmological. The antigen and
 CC its encoded gene are for producing drug compositions in treating
 CC allergic diseases e.g. atopic dermatitis, asthma, allergic rhinitis
 CC and allergic conjunctivitis. The drug compositions are made from the
 CC protein that can inhibit production of interleukin-4, ameliorate
 CC Th2-type cytokine-predominate immune state and relieve
 CC symptoms of allergic diseases such as IGE production, histamine
 CC release and eosinophilic infiltration. The present sequence
 CC represents the mycobacterium-originated alpha antigen. This
 CC sequence is not encoded by the sequence featured in ABQ79295 despite
 CC being said to do so in the specification.
 XX Sequence 285 AA;
 SQ Query Match 82.3%; Score 1437; DB 23; Length 285;
 Best Local Similarity 90.5%; Pred. No. 6.8e-117;
 Matches 257; Conservative 18; Mismatches 9; Indels 0; Gaps 0;
 QY 41 FSRPGLPVEYLQVPSMGRDIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFW 100
 DB 1 FSRPGLPVEYLQVPSMGRDIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFW 60
 QY 101 YQSGLSIVMPVGGQSFYSQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFW 160
 DB 61 YQSGLSIVMPVGGQSFYSQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFW 120
 QY 161 AATGLSMAGSSAMILAAHPQFIYAGSLALDPQSGMGPSLIGLWAGDAGGYKAADMW 220
 DB 121 AAVGISWAGSSALILSVYHPQFIYAGSLALMDPQSGMGPSLIGLWAGDAGGYKAADMW 180
 QY 221 GPSSDPAWERNDDTQIIPKLVANNTLWYVCGNGTPELNGGNIIPAEFLFNFRSSNLK 280
 DB 181 GPSSDPAWERNDDTQIIPKLVANNTLWYVCGNGTPELNGGNIIPAEFLFNFRSSNLK 240
 QY 281 QDAYNAGGHNVAFFNPNGTHSWEYWGALNMGDLOSSLGA 324
 DB 241 QDAYNAGGHNVAFFNPNGTHSWEYWGALNMGDLOSSLGA 284

RESULT 22
 ID ABB83662 standard; Protein; 337 AA.
 XX ABB83662;
 AC ABB83662;

XX DT 31-MAR-2003 (first entry)

XX DE Mycobacterium tuberculosis 32kDa protein.

XX KW Antibacterial; tuberculosis; vaccine; gene therapy; 32kDa protein.

XX OS Mycobacterium tuberculosis.

XX PN US2002131975-A1.

XX PD 19-SEP-2002.

XX PF 14-SEP-2001; 2001US-0953510.

XX PR 23-MAY-1996; 96US-0652842.

XX PR 23-NOV-1993; 93US-0156358.

XX PR 12-AUG-1994; 94US-0289667.

XX PR 23-MAY-1995; 95US-0447398.

XX PR 31-OCT-1995; 95US-0551149.

XX PR 06-DEC-1996; 96US-0568357.

XX PR 21-SEP-1998; 98US-0157689.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Horwitz MA, Harth G;

XX DR WPI; 2003-174073/17.

XX DR N-PSDB; ABX75989.

XX PT New vaccine, useful for promoting an immune response against infectious pathogens of the genus Mycobacterium in a mammalian host

XX PS Example 1; Page 15-16; 82pp; English.

XX CC The invention describes a vaccine for promoting an immune response, in a mammalian host, against infectious pathogens of the genus Mycobacterium, comprising at least 1 immunodominant epitope of at least one majority abundant extracellular product comprising Mycobacterium tuberculosis 110, 80, 71, 58, 45, 32A, 32B, 30, 24, 23.5, 23, 16, 14, 12 kD protein or their analogues, homologues or subunits.

XX CC The proteins and polypeptides of the invention are useful in gene therapy and treatment of diseases caused by Mycobacterium such as tuberculosis. This is the amino acid sequence of the Mycobacterium tuberculosis 32kDa protein.

XX SQ Sequence 337 AA;

Query Match 80.6%; Score 1407; DB 24; Length 337;

Best Local Similarity 78.9%; Pred. No. 3.5e-114; Mismatches 38; Indels 0; Gaps 0;

Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

QY 8 IRWGRRLMIGTAAAVVLFGLVGLAGGATAGAFSRPGLPVEYLQVPSMGRDIKVQFQ 67

DB 11 VTGMSRRLVVGGAALVSLGVAGVGTATACAFSRPGLPVEYLQVPSMGRDIKVQFQ 70

QY 68 SCGNNSPAVLLDGLRAQDYGNDINTPAFWEYVQSGLSIWMPVGGQSSFYSDWYSAC 127

DB 71 SGGANSPLALYLLDGLRAQDDFGSDINTPAFEWYDQGLSVMPVGGQSSFYSDWYQAC 130

QY 128 GKAGQCTYKWEFTLSELPWLSANRAVKPTGSAAGLSMAGSSAMILAAYHPQOFIVAG 187

DB 131 GKAGQCTYKWEFTLSELPWLSANRAVKPTGSAAGLSMAGSSAMILAAYHPQOFIVAG 190

QY 188 SLISALLDPQSGMPSLIGLAMDAGGYKAADWMPSSDPAWRNPTQOIPKLNVNREL 247

DB 191 AMSGLLDPSQMGPTLIGLAMDAGGYKASDMWGPKEPDPAWQRNDPLLVNGLIANTRV 250

QY 248 WYCGNGTNEIGGANIPAEFLNFVRSNLKFDQAYNAAGGNVAFNPPNGTHSEWY 307

DB 251 WYCGNGKPSDLGGNNLPKAFLEGFVRSNLIKFDQAYNAAGGGHGVDFPDGTHSEWY 310

QY 308 GAQLNAMKGLQSSIGA 324

Db 311 GAQLNAMKPDQLQALGA 327

RESULT 23

AAW18165

ID AAW18165 standard; Protein; 338 AA.

XX AC AAW18165;

XX DT 13-AUG-1997 (first entry)

XX DE Mycobacterium tuberculosis extracellular 32AKD protein.

XX KW Vaccine; vaccinating agent; M.tuberculosis; pathogen; bacteria; virus; fungus; protozoan; HIV.

XX OS Mycobacterium tuberculosis.

XX PH Key Location/Qualifiers

XX FT Protein 44..338

XX FT /label= Mature

XX PN W09637219-A1.

XX PD 28-NOV-1996.

XX PF 23-MAY-1996; 96WO-US07781.

XX PR 06-DEC-1995; 95US-0568357.

XX PR 23-MAY-1995; 95US-0447398.

XX PR 20-OCT-1995; 95US-0545926.

XX PR 31-OCT-1995; 95US-0551149.

XX PA (REGC) UNIV CALIFORNIA.

XX PI Harth G, Horwitz MA;

XX DR WPI; 1997-020936/02.

XX DR N-PSDB; AAT71599.

XX PT Vaccines derived from M.tuberculosis major abundant extracellular proteins - are easy to prepare and less toxic than conventional killed or attenuated vaccines, useful for protecting against or treating Mycobacterial infections

XX PS Claim 1; Page 36-38; 193pp; English.

XX CC A vaccinating agent for promoting an immune response in a mammal against Mycobacterium pathogens comprises at least one majorly abundant extracellular protein, i.e. the M.tuberculosis 110, 80, 71, 58, 45, 32A, 32B, 30, 24, 23.5, 23, 16, 14 or 12 kD proteins, or their analogues, homologues and subunits. The present sequence represents the 32A kD protein. The vaccinating agents are used to protect against (or to treat existing) infections by Mycobacterium (especially M. tuberculosis) while the epitopes can also be used to detect presence of an immune response to a Mycobacterium pathogen. The vectors, containing the DNA for the extracellular proteins, are used to transform cells for production of recombinant DNA molecules. More generally the DNA from other pathogens can be used in vaccines, e.g. against other bacteria, viruses, fungi and protozoa. Since different combinations of DNA can be used, a wide range of effective compositions can be produced. They generate a response against the antigens most often found on infected cells during the infection, regardless of the strength or specificity of the immune response. The vaccines are easy to produce and less toxic than known killed or attenuated vaccines, so can be given to immunocompromised subjects, e.g. those with HIV infection.

XX SQ Sequence 338 AA;

Query Match 80.6%; Score 1407; DB 18; Length 338;

Best Local Similarity 78.9%; Pred. No. 3.5e-114; Mismatches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

XX PI Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
XX DR WPI; 1999-430163/36.
XX PT Enhancing immune response to an antigen
XX PS Example 11; Page 155-156; 243pp; English.
XX CC The invention provides heat-killed Mycobacterium vaccae, or recombinant
CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
CC T cells and natural killer cells, to stimulate the production of
CC cytokines, to enhance the expression of co-stimulatory molecules on
CC dendritic cells and monocytes, and to enhance dendritic cell maturation
CC and function. The proteins can be expressed by standard recombinant
CC methodology. Pharmaceutical compositions comprising the proteins or
CC nucleic acid sequences encoding the proteins can be used for the
CC treatment, prevention, and detection of disorders including infectious
CC diseases, immune disorders and cancer. In particular, the compounds and
CC methods are used for treatment of diseases of the respiratory system,
CC such as mycobacterial infections, asthma, allergies, tuberculosis,
CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
CC carcinoma and melanoma.
XX SQ Sequence 338 AA;
Query Match 80.6%; Score 1407; DB 20; Length 338;
Best Local Similarity 78.9%; Pred. No. 3.5e-114;
Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;
QY 8 IRWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGRDIKVQFQ 67
DB 11 VTGMSRLVVGAVGAALVSLVGLVAVGTATAGAFSRPGLPVEYLQVPSMGRDIKVQFQ 70
QY 68 SGGNNSPAILLDGLRAQDDYNGWDINTPAFEWYQSGLSIYMPVGGSSFYSDWYSPAC 127
DB 71 SGANSALYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIYMPVGGSSFYSDWYSPAC 130
QY 128 GKAGCQTYKWTFLTSELFWLSANRAVKPTGSAALGLSMAGSSAMLAAYHPQOFIYAG 187
DB 131 GKAGCQTYKWTFLTSELFWLSANRAVKPTGSAALGLSMAGSSAMLAAYHPQOFIYAG 190
QY 188 SILSALLDPQGMGSPSLIGLWAGDAGYKAADMMWGPSSDPAWERNDDTQOIPKLVANTRL 247
DB 191 AMSGLLDPSQAMGPTLLIGLWAGDAGYKASDMWGPKEDEPAWQRNDPDLNVGKLIANNTRV 250
QY 248 WYTCNGTNPNEGNIPIAEFLNFVRSSNLFQDAYNAGGHNAPFNPPNGTHSWEYW 307
DB 251 WYTCNGKPSDLGGNNLPKFLFEGFVRTSNIKFQDAYNAGGHNAPFNPPNGTHSWEYW 310
QY 308 GAQLNAMKGDLOSSLGA 324
DB 311 GAQLNAMKPDQLRALGA 327
RESULT 26
AA14854
ID AA14854 standard; protein; 338 AA.
XX AC AA14854;
XX DT 25-OCT-1999 (first entry)
XX DE Antigen 85A protein from M. bovis.
XX KW Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
KW dendritic cell maturation; infectious disease; immune disorder; cancer;
KW respiratory system; mycobacterial infection; allergy; tuberculosis;
KW leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
KW dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
KW squamous cell carcinoma; melanoma.

XX OS Mycobacterium bovis.
XX PN WO9932634-A2.
XX XX 01-JUL-1999.
XX PF 23-DEC-1998; 98WO-NZ00189.
XX PR 04-DEC-1998; 98US-0205426.
PR 23-DEC-1997; 97US-0996624.
PR 23-DEC-1997; 97US-0997080.
PR 23-DEC-1997; 97US-0997362.
PR 11-JUN-1998; 98US-0095855.
PR 17-SEP-1998; 98US-0156181.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX PA Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
XX PI WPI; 1999-430163/36.
XX DR Enhancing immune response to an antigen
XX PT Example 11; Page 157-158; 243pp; English.
XX PS The invention provides heat-killed Mycobacterium vaccae, or recombinant
XX CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
CC T cells and natural killer cells, to stimulate the production of
CC cytokines, to enhance the expression of co-stimulatory molecules on
CC dendritic cells and monocytes, and to enhance dendritic cell maturation
CC and function. The proteins can be expressed by standard recombinant
CC methodology. Pharmaceutical compositions comprising the proteins or
CC nucleic acid sequences encoding the proteins can be used for the
CC treatment, prevention, and detection of disorders including infectious
CC diseases, immune disorders and cancer. In particular, the compounds and
CC methods are used for treatment of diseases of the respiratory system,
CC such as mycobacterial infections, asthma, allergies, tuberculosis,
CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
CC carcinoma and melanoma.
XX SQ Sequence 338 AA;
Query Match 80.6%; Score 1407; DB 20; Length 338;
Best Local Similarity 78.9%; Pred. No. 3.5e-114;
Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;
QY 8 IRWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGRDIKVQFQ 67
DB 11 VTGMSRLVVGAVGAALVSLVGLVAVGTATAGAFSRPGLPVEYLQVPSMGRDIKVQFQ 70
QY 68 SGGNNSPAILLDGLRAQDDYNGWDINTPAFEWYQSGLSIYMPVGGSSFYSDWYSPAC 127
DB 71 SGANSALYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIYMPVGGSSFYSDWYSPAC 130
QY 128 GKAGCQTYKWTFLTSELFWLSANRAVKPTGSAALGLSMAGSSAMLAAYHPQOFIYAG 187
DB 131 GKAGCQTYKWTFLTSELFWLSANRAVKPTGSAALGLSMAGSSAMLAAYHPQOFIYAG 190
QY 188 SILSALLDPQGMGSPSLIGLWAGDAGYKAADMMWGPSSDPAWERNDDTQOIPKLVANTRL 247
DB 191 AMSGLLDPSQAMGPTLLIGLWAGDAGYKASDMWGPKEDEPAWQRNDPDLNVGKLIANNTRV 250
QY 248 WYTCNGTNPNEGNIPIAEFLNFVRSSNLFQDAYNAGGHNAPFNPPNGTHSWEYW 307
DB 251 WYTCNGKPSDLGGNNLPKFLFEGFVRTSNIKFQDAYNAGGHNAPFNPPNGTHSWEYW 310
QY 308 GAQLNAMKGDLOSSLGA 324
DB 311 GAQLNAMKPDQLRALGA 327

QY	128	GKAGCQTYKWETFLITSELP	OWLSANRAVKPTGSAATLMSAGSSAMILAAAYHPQQFIYAG	187
Db	131	GKAGCQTYKWETFLITSELP	PCWLQANRHVKTGSAAVLGLSVAASSALTALAIYHPQQFVIYAG	190
QY	188	SLSALLDPSQMGPSLIGLAMGDAGGYKAADMMGPSSDP	PAWERNDFTQIIPKLVANNTRL	247
Db	191	AMSGLDPSQAMGFTLIGLAMGDAGGYKASDMWGPKE	DPAWRNDPFLNNGVGLIANNTLV	250
QY	248	WVYCNGTNPENELGANIPAEFLNFVRSSNLKFQDAYNAAGHNAVFNFPNGTHSWEYW	307	
Db	251	WVYCNGKPSDLGGNNLPARLEFGFVRTSNIKRQDAYNAGGGHGVDFFDSTHSWEYW	310	
QY	308	GAQLNAMKGDLQSSSLGA	324	
Db	311	GAQLNAMKPDQRALGA	327	
RESULT 28				
ID	AAB82789			
XX	AAAB82789 standard; Protein; 338 AA.			
AC	AAAB82789;			
DT	29-OCT-2001 (first entry)			
XX	Mycobacterium tuberculosis antigen 85A.			
DE	Mycobacterial antigen 85A; Ag85A; infection; vaccine; diagnosis; therapy.			
KW	Mycobacterium tuberculosis.			
OS	Mycobacterium tuberculosis.			
XX	Key	Location/Qualifiers		
FH	Peptide	48..56		
FT	Peptide	/note= "epitope"		
FT	Peptide	242..250		
FT	Peptide	/note= "epitope"		
XX	WO200158461-A1.			
PN	16-AUG-2001.			
XX	12-FEB-2001; 2001WO-GB00561.			
PR	10-FEB-2000; 2000GB-0003082.			
XX	(GLAX) GLAXO GROUP LTD.			
PI	Dockrell HM, Smith SM, Brookes R;			
DR	WPI; 2001-536505/59.			
DR	N-PSDB; AAH26485.			
XX	Use of polypeptides comprising groups of mycobacterial antigen 85A protein and expression vectors comprising polynucleotides encoding the polypeptide for vaccination against Mycobacterium tuberculosis -			
PS	Disclosure; Page 42-43; 49pp; English.			
CC	The present sequence is that of Mycobacterium tuberculosis antigen 85A (Ag85A). Epitopes of Ag85A, such as those including amino acids 48-56 or 252-250 of the present sequence (see AAB82787-88), have been identified which cause the generation of a strong CD8 T-cell response in humans. CD8 T-cells recognising such epitopes are able to lyse macrophages infected with live M. tuberculosis. A polypeptide which comprises an epitope structure of Ag85A, or an expression vector comprising a polynucleotide encoding such a polypeptide, is used in the manufacture of a medicament for vaccinating prophylactically or therapeutically against infection by a mycobacterium stimulating a CD8 T-cell response. The polypeptide or expression vector may be within an antigen-presenting cell. A claimed vaccine composition comprises the polypeptide or			

CC expression vector and an adjuvant or delivery system capable of
CC stimulating a CD8 T-cell response. Also claimed are methods of
CC detecting in a population of T-cells the presence or absence of CD8
CC T-cells that recognise the epitope sequence, a method of diagnosing
CC mycobacterial infection or of testing the effectiveness of a
CC vaccination, and a method of treating a mycobacterium infection by
CC administering T-cells capable of recognising the epitope.

XX SQ Sequence 338 AA;

Query Match 80.6%; Score 1407; DB 22; Length 338;
Best Local Similarity 78.9%; Pred. No. 3.5e-114;
Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

QY 8 IRAGRLMIGTAAAVVLGLVGLAGGAATAGAFSPRLPVEYLQVPSMGRDIKVFQF 67
DB 11 VTGMSRLVVGAVGAAALVGLVAVGTATAGAFSPRLPVEYLQVPSMGRDIKVFQF 70
QY 68 SGGNNSPAVILLDLGRAQDDYNGWDINTPAPEWYQSGLSIVMPVGGSSFYSDWYSPAC 127
DB 71 SGGANSFALYLLDLGRAQDDFSGWDINTPAPEWYDQSGLSVMPVGGSSFYSDWYQAC 130
QY 128 GKACQYKWTFTLTSELPOWLSANRAVKPTGSAAGLSMAGSAMILLAAVHPQOFIYAG 187
DB 131 GKACQYKWTFTLTSELPGWLQANRHVKPTGSAVGLSMAASSALTILAIHPQOFVYAG 190
QY 188 SLSALLDPSQMGPSLIGLAMDAGGYKAAADWMPGSSDPAPERNDPTQOIPKLIVANNTRL 247
DB 191 AMSGLDPSQMGPTLIGLAMDAGGYKASDMWGPKEPAPQWRNDPLINVGKLIANNTRV 250
QY 248 WYCGNGTNPGLGGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVENFPNGTHSWEYW 307
DB 251 WYCGNGKPSDLGGNNLPKAFLEGFVRTSNIKFQDAYNAGGCHNGVDFPDGTHSWEYW 310
QY 308 GAQLNAMKGDLOSLIGA 324
DB 311 GAQLNAMKPDLOALGA 327

RESULT 29
ABB73458
ID ABB73458 standard; Protein; 338 AA.
XX AC ABB73458;
XX DT 08-APR-2002 (first entry)
XX DE M tuberculosis 85A protein SEQ ID NO: 32.

XX skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic;
KW Th2 immune response; immunomodulatory.

XX Mycobacterium tuberculosis.
XX US6328978-B1.
XX 11-DEC-2001.
XX 02-JUN-1999; 99US-0324542.
XX 23-DEC-1997; 97US-0997080.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Watson JD, Tan PLJ, Prestidge R;
XX WPI; 2002-138361/18.

XX Inhibiting skin inflammation associated with skin disorder e.g.
PT psoriasis, by administering composition comprising delipidated and
PT deglycolipidated Mycobacterium vaccae cells or Mycobacterium vaccae

PT culture filtrate -

XX Example 5; Column 59-62; 116pp; English.

XX The present invention relates to a method of inhibiting skin inflammation
CC associated with a skin disorder selected from psoriasis, atopic
CC dermatitis and allergic contact dermatitis, which involves administering
CC a composition containing delipidated and deglycolipidated Mycobacterium
CC vaccae cells or M. vaccae culture filtrate. The skin disorder to be
CC treated may also include alopecia areata, and skin cancers such as basal
CC cell carcinoma, squamous cell carcinoma and melanoma. The composition
CC acts by inhibiting the Th2 immune response. The present sequence is a
CC protein described in the exemplification of the invention.

XX SQ Sequence 338 AA;

Query Match 80.6%; Score 1407; DB 23; Length 338;
Best Local Similarity 78.9%; Pred. No. 3.5e-114;
Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

QY 8 IRAGRLMIGTAAAVVLGLVGLAGGAATAGAFSPRLPVEYLQVPSMGRDIKVFQF 67
DB 11 VTGMSRLVVGAVGAAALVGLVAVGTATAGAFSPRLPVEYLQVPSMGRDIKVFQF 70
QY 68 SGGNNSPAVILLDLGRAQDDYNGWDINTPAPEWYQSGLSIVMPVGGSSFYSDWYSPAC 127
DB 71 SGGANSFALYLLDLGRAQDDFSGWDINTPAPEWYDQSGLSVMPVGGSSFYSDWYQAC 130
QY 128 GKACQYKWTFTLTSELPOWLSANRAVKPTGSAAGLSMAGSAMILLAAVHPQOFIYAG 187
DB 131 GKACQYKWTFTLTSELPGWLQANRHVKPTGSAVGLSMAASSALTILAIHPQOFVYAG 190
QY 188 SLSALLDPSQMGPSLIGLAMDAGGYKAAADWMPGSSDPAPERNDPTQOIPKLIVANNTRL 247
DB 191 AMSGLDPSQMGPTLIGLAMDAGGYKASDMWGPKEPAPQWRNDPLINVGKLIANNTRV 250
QY 248 WYCGNGTNPGLGGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVENFPNGTHSWEYW 307
DB 251 WYCGNGKPSDLGGNNLPKAFLEGFVRTSNIKFQDAYNAGGCHNGVDFPDGTHSWEYW 310
QY 308 GAQLNAMKGDLOSLIGA 324
DB 311 GAQLNAMKPDLOALGA 327

RESULT 30
ABB73460
ID ABB73460 standard; Protein; 338 AA.
XX AC ABB73460;
XX DT 08-APR-2002 (first entry)
XX DE M bovis 85A protein SEQ ID NO: 34.

XX skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
KW antipsoriatic; dermatological; antiinflammatory; antiallergic;
KW Th2 immune response; immunomodulatory.

XX Mycobacterium bovis.
XX US6328978-B1.
XX 11-DEC-2001.
XX 02-JUN-1999; 99US-0324542.
XX 23-DEC-1997; 97US-0997080.
XX (GENE-) GENESIS RES & DEV CORP LTD.
XX Watson JD, Tan PLJ, Prestidge R;
XX

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:35:17 ; Search time 21.6667 Seconds
(without alignments)
1442.530 Million cell updates/sec

Title: US-09-805-427A-2
Perfect score: 1745
Sequence: 1 MTDVSRKIRAWGRRLMIGTA.....YKGAQLNAMKGLQSSLCAG 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 9616862 residues

Total number of hits satisfying chosen parameters: 283102

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database :

PIR_76.*

1: PIR1.*

2: PIR2.*

3: PIR3.*

4: PIR4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1745	100.0	325	2 C70516	probable fbpB prot
2	1739	99.7	325	2 S29663	antigen 85-B precu
3	1706	97.8	323	2 A32348	alpha-antigen B pr
4	1609	92.2	325	2 A37185	alpha-antigen prec
5	1539	88.2	330	2 JN0897	alpha-antigen prec
6	1519	87.0	330	2 S32773	alpha-antigen - My
7	1455	83.4	327	2 G87162	antigen 85A, mycol
8	1454	83.3	327	2 S34434	antigen 85-B - Myc
9	1428	81.8	327	2 S20038	fibronectin-bindin
10	1407	80.6	338	2 S10326	alpha-antigen A, e
11	1407	80.6	338	2 H70887	32K antigen fbpA p
12	1338.5	76.7	330	2 A86921	antigen 85A, mycol
13	1198	68.7	340	2 D70615	antigen fbpC2 - My
14	1170	67.0	333	2 S32114	85C protein - Myco
15	775	44.4	188	2 S32107	fbpC1 protein - Myco
16	519	29.7	299	2 G70887	antigen 85C, mycol
17	498.5	28.6	301	2 B68921	cspl protein - Cor
18	474	27.2	657	2 S25184	MPT51 protein - My
19	367.5	21.1	220	2 S32111	antigen 85-C - Myc
20	257	14.7	57	2 B43603	hypothetical prote
21	233.5	13.4	342	2 T29115	hypothetical prote
22	146.5	8.4	456	2 D70772	tributyrin ester
23	143	8.2	259	2 E95071	tributyrin ester
24	143	8.2	259	2 C97939	acetyltransferase hom
25	134	7.7	252	2 AB1748	lipase [imported]
26	133	7.6	252	2 C11378	hypothetical prote
27	132	7.6	258	2 C86846	hypothetical prote
28	129.5	7.4	76	2 A70716	alpha-antigen C, e
29	128	7.3	26	2 C41499	

ALIGNMENTS

RESULT 1

C70516

probable fbpB protein - Mycobacterium tuberculosis (strain H37RV)

N:Alternate names: antigen 85-B precursor

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: C70516; S29664

probable membrane
antigen A 31K comp
S-formylglutathion
probable esterase
esterase/lipase (i
tributyrin ester
prolyl oligopeptid
probable esterase
transferred entry
esterase sll0992 -
prolyl endopeptida
poly(3-hydroxybuty
hypothetical prote
probable transmem
hypothetical prote
carboxylesterase (i
methylesterase
hypothetical prote
probable esterase
carboxylesterase (i
transforming prote
hypothetical prote
hypothetical prote
regulatory protein
esterase [imported
hypothetical prote
hypothetical prote
YalM protein - Esc
proteoglycan core
aggreccan - bovine
cellulase (EC 3.2.
probable transmem
hypothetical prote
probable pyrophosph
hypothetical prote
porin P precursor,
outer membrane por
poly(3-hydroxybuty
probable esterase
probable lpqr prot
conserved hypotet
lep1170 F3 112 pr
hypothetical prote
hypothetical prote
beta-fructofuranos
hypothetical prote
hypothetical prote
conserved hypotet
conserved hypotet
period clock prote
branched-chain ami
diene lactone hydro
hypothetical prote
pectate lyase (EC
pyruvate, phosphat
conserved hypotet
cellulose 1,4-beta
hypothetical prote
probable photosyst
and cellulose-bind

7.2 377 2 T36322
6.6 26 2 A44892
6.5 282 2 A12156
6.4 284 2 T00809
6.4 322 2 A28222
6.4 322 2 E97600
6.4 618 2 D87651
6.2 363 2 T36408
6.2 827 2 J04900
6.2 276 2 S75249
6.2 689 2 A21222
6.1 414 2 S39530
6.0 293 2 C90032
6.0 415 2 T36279
6.0 430 2 B81633
6.0 280 2 A10192
6.0 297 2 A23543
6.0 963 2 T40290
5.9 277 2 B90680
5.9 277 2 F85530
5.8 1106 1 TVHUGL
5.8 328 2 C87451
5.8 426 2 S28812
5.7 555 1 RGASWA
5.6 277 2 B87561
5.6 472 2 C75505
5.6 1501 2 T45623
5.6 277 2 C64763
5.6 1340 2 A39808
5.6 2327 2 T42630
5.6 264 2 JU0328
5.6 344 2 T36160
5.6 359 2 H70876
5.6 794 2 T36668
5.5 631 2 B83404
5.5 440 2 S11793
5.5 440 2 F83235
5.5 488 2 A32235
5.4 285 2 AB0782
5.4 452 2 F70552
5.4 358 2 B87063
5.4 391 2 S72717
5.4 312 2 A70708
5.4 489 2 B70619
5.4 567 2 AF0308
5.4 593 2 T01575
5.4 392 2 S72984
5.4 420 2 AG1859
5.4 473 2 E86955
5.4 884 2 E75489
5.4 1122 2 B26427
5.3 424 2 G75579
5.3 323 2 G87358
5.3 536 2 A99283
5.3 677 2 JC7303
5.3 601 2 E87028
5.3 786 2 E87565
5.3 1230 2 S47466
5.2 299 2 S50803
5.2 403 2 T51828
5.2 537 2 B97013

ALIGNMENTS

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, K.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.R.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: C70516
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-325 <COL>
A:Cross-references: GB:297193; GB:AL123456; NID:g3261816; PIDN:CAB10044.1; PID:g2225974
A:Experimental source: strain H37Rv
R:de Wit, L.; Palou, M.; Content, J.
submitted to the EMBL Data Library, December 1991
A:Reference number: S29663
A:Accession: S29664
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <WIT>
A:Cross-references: EMBL:X62398; NID:g44563; PIDN:CAA44269.1; PID:g44564
C:Genetics: fbpB
A:Gene: fbpB
C:Superfamily: Mycobacterium avium alpha-antigen

Query Match 100.0%; Score 1745; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.4e-126; Mismatches 0; Indels 0; Gaps 0;
Matches 325; Conservative 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
Db 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
QY 121 DWYSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
Db 121 DWYSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
QY 181 QOFIYAGLSALLDPSQMGPSLIGLAMDAGGYKAADMWGPSSDPDAWERNPTQOIPLK 240
Db 181 QOFIYAGLSALLDPSQMGPSLIGLAMDAGGYKAADMWGPSSDPDAWERNPTQOIPLK 240
QY 241 VANNTLWYCGNGTPELGGANI PAEFLNFVRSSNLKFQDAYNAAGGHNAVFNPFG 300
Db 241 VANNTLWYCGNGTPELGGANI PAEFLNFVRSSNLKFQDAYNAAGGHNAVFNPFG 300
QY 301 THSWEYWGQALNAMKGDLOSSLGAG 325
Db 301 THSWEYWGQALNAMKGDLOSSLGAG 325

RESULT 2

S29663
antigen 85-B precursor - Mycobacterium bovis
C:Species: Mycobacterium bovis
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-May-2000
C:Accession: S29663
R:de Wit, L.; Palou, M.; Content, J.
submitted to the EMBL Data Library, December 1991
A:Reference number: S29663
A:Accession: S29663
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-325 <WIT>
A:Cross-references: EMBL:X62397; NID:g44167; PIDN:CAA44268.1; PID:g44168
C:Superfamily: Mycobacterium avium alpha-antigen

Query Match 99.7%; Score 1739; DB 2; Length 325;
Best Local Similarity 99.7%; Pred. No. 7e-126;
Matches 324; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
Db 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
QY 121 DWYSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
Db 121 DWYSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
QY 181 QOFIYAGLSALLDPSQMGPSLIGLAMDAGGYKAADMWGPSSDPDAWERNPTQOIPLK 240
Db 181 QOFIYAGLSALLDPSQMGPSLIGLAMDAGGYKAADMWGPSSDPDAWERNPTQOIPLK 240
QY 241 VANNTLWYCGNGTPELGGANI PAEFLNFVRSSNLKFQDAYNAAGGHNAVFNPFG 300
Db 241 VANNTLWYCGNGTPELGGANI PAEFLNFVRSSNLKFQDAYNAAGGHNAVFNPFG 300
QY 301 THSWEYWGQALNAMKGDLOSSLGAG 325
Db 301 THSWEYWGQALNAMKGDLOSSLGAG 325

RESULT 3

A32348
alpha-antigen B precursor, extracellular - Mycobacterium bovis
N:Alternate names: 32k antigen
C:Species: Mycobacterium bovis
C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 26-May-2000
C:Accession: A32348; A41499; A60278
R:Matsuo, K.; Yamaguchi, R.; Yamazaki, A.; Tasaka, H.; Yamada, T.
J. Bacteriol. 170, 3847-3854, 1988
A:Title: Cloning and expression of the Mycobacterium bovis BCG gene for extracellular
A:Reference number: A32348; MUID:88314872; PMID:2842287
A:Accession: A32348
A:Molecule type: DNA
A:Residues: 1-323 <MA>
A:Cross-references: GB:M21839; NID:g149935; PIDN:AAA25359.1; PID:g149936
R:Wiker, H.G.; Sletten, K.; Nagai, S.; Harboe, M.
Infect. Immun. 58, 272-274, 1990
A:Title: Evidence for three separate genes encoding the proteins of the mycobacterial
A:Reference number: A41499; MUID:90093478; PMID:2403534
A:Accession: A41499
A:Molecule type: protein
A:Residues: 41-85 <WIK>
R:Fifis, T.; Costopoulos, C.; Radford, A.J.; Bacic, A.; Wood, P.R.
Infect. Immun. 59, 800-807, 1991
A:Title: Purification and characterization of major antigens from a Mycobacterium bovis
A:Reference number: A60278; MUID:91147217; PMID:1900061
A:Accession: A60278
A:Molecule type: protein
A:Residues: 41-59 <PIF>
C:Superfamily: Mycobacterium avium alpha-antigen
F:1-40/Domain: signal sequence #status predicted <SIG>
F:41-323/Product: alpha antigen, extracellular #status experimental <MAT>

Query Match 97.8%; Score 1706; DB 2; Length 323;
Best Local Similarity 98.5%; Pred. No. 2.3e-123;
Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
Db 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
QY 121 DWYSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
Db 121 DWYSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180

Db 121 DWYSPACGKAGCQTYKWTETLTSELPAQLSANRAVKPTGSAALGLSMAGSSAMILAAVHP 180
 QY 181 QOPIYAGSLALLDPSQGMPSLIGLMDGAGGYKAADMWGPSSDPAWERNPTQOIPKL 240
 Db 181 QOPIYAGSLALLDPSQGMPSLIGLMDGAGGYKAADMWGPSSDPAWERNPTQOIPKL 238
 QY 241 VANNTLWYCGNGTPELGGANIPAEFLNFVRSNLKFDAYNAAGGHNAVFPPNG 300
 Db 239 VANNTLWYCGNGTPELGGANIPAEFLNFVRSNLKFDAYNAAGGHNAVFPPNG 298
 QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
 Db 299 THSWEYGAQLNAMKGDLOSSLGAG 323

RESULT 4

A37185

alpha-antigen precursor - Mycobacterium kansasii
 C:Species: Mycobacterium kansasii

C:Date: 31-Jan-1992 #sequence_revision 31-Jan-1992 #text_change 26-May-2000

C:Accession: A37185

R;Matsuo, K.; Yamaguchi, R.; Yamazaki, A.; Tasaka, H.; Terasaka, K.; Yamada, T.
 Infect. Immun. 58, 550-556, 1990

A;Title: Cloning and expression of the gene for the cross-reactive alpha antigen of Myco
 A;Reference number: A37185; MUID:90129315; PMID:2404875

A;Accession: A37185

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-325 <MAT>

A;Cross-references: GB:M53897

C;Superfamily: Mycobacterium avium alpha-antigen

Query Match 92.2%; Score 1609; DB 2; Length 325;
 Best Local Similarity 89.8%; Pred. No. 6.4e-116;
 Matches 291; Conservative 20; Mismatches 13; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
 Db 1 MTDVSGKIRAWGRRLLVGAATAAALPGLVLGAGGAATAGAFSRPGLPVEYLQVPSAAGR 60

QY 61 DIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSLGSLVMPVGGQSSFFS 120
 Db 61 SIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSLGSLVMPVGGQSSFFS 120

QY 121 DWYSPACGKAGCQTYKWTETLTSELPAQLSANRAVKPTGSAALGLSMAGSSAMILAAVHP 180
 Db 121 DWYSPACGKAGCQTYKWTETLTSELPAQLSANRAVKPTGSAALGLSMAGSSAMILAAVHP 180

QY 181 QOPIYAGSLALLDPSQGMPSLIGLMDGAGGYKAADMWGPSSDPAWERNPTQOIPKL 240
 Db 181 QOPIYAGSLALLDPSQGMPSLIGLMDGAGGYKAADMWGPSSDPAWERNPTQOIPKL 240

QY 241 VANNTLWYCGNGTPELGGANIPAEFLNFVRSNLKFDAYNAAGGHNAVFPPNG 300
 Db 241 VANNTLWYCGNGTPELGGANIPAEFLNFVRSNLKFDAYNAAGGHNAVFPPNG 300

QY 301 THSWEYGAQLNAMKGDLOSSLGAG 324
 Db 301 THSWEYGAQLNAMKGDLOSSLGAG 324

RESULT 5

JN0897

alpha-antigen precursor - Mycobacterium intracellulare
 N;Alternate names: A985B

C:Species: Mycobacterium intracellulare

C:Date: 19-May-1994 #sequence_revision 19-May-1994 #text_change 26-May-2000

C:Accession: JN0897

R;Kitaura, H.; Ohara, N.; Matsuo, T.; Tasaka, H.; Kobayashi, K.; Yamada, T.

Biochem. Biophys. Res. Commun. 196, 1466-1473, 1993

A;Title: Cloning, sequencing and expression of the gene for alpha antigen from Mycobacte

A;Reference number: JN0897; MUID:94071912; PMID:8250904

A;Accession: JN0897

A;Molecule type: DNA

A;Residues: 1-330 <KIT>

C;Comment: This protein is one of antigen 85.

C;Superfamily: Mycobacterium avium alpha-antigen

F;1-40/Domain: signal sequence #status predicted <SIG>

F;41-330/Product: alpha-antigen #status predicted <PAA>

Query Match

Best Local Similarity 88.2%; Score 1539; DB 2; Length 330;

Matches 278; Conservative 23; Mismatches 23; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
 Db 1 MTDLSEKVRWGRRLVGAATAATLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60

QY 61 DIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSLGSLVMPVGGQSSFFS 120
 Db 61 DIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSLGSLVMPVGGQSSFFS 120

QY 121 DWYSPACGKAGCQTYKWTETLTSELPAQLSANRAVKPTGSAALGLSMAGSSAMILAAVHP 180
 Db 121 DWYSPACGKAGCQTYKWTETLTSELPAQLSANRAVKPTGSAALGLSMAGSSAMILAAVHP 180

QY 181 QOPIYAGSLALLDPSQGMPSLIGLMDGAGGYKAADMWGPSSDPAWERNPTQOIPKL 240
 Db 181 QOPIYAGSLALLDPSQGMPSLIGLMDGAGGYKAADMWGPSSDPAWERNPTQOIPKL 240

QY 241 VANNTLWYCGNGTPELGGANIPAEFLNFVRSNLKFDAYNAAGGHNAVFPPNG 300
 Db 241 VANNTLWYCGNGTPELGGANIPAEFLNFVRSNLKFDAYNAAGGHNAVFPPNG 300

QY 301 THSWEYGAQLNAMKGDLOSSLGAG 324
 Db 301 THSWEYGAQLNAMKGDLOSSLGAG 324

RESULT 6

S32773

alpha-antigen - Mycobacterium avium

C:Species: Mycobacterium avium

C;Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 26-May-2000

C;Accession: S32773

R;Ohara, N.; Matsuo, K.; Yamaguchi, R.; Yamazaki, A.; Tasaka, H.; Yamada, T.

submitted to the EMBL Data Library, December 1991

A;Description: Nucleotide sequence of the alpha-antigen gene of Mycobacterium avium.

A;Reference number: S32773

A;Accession: S32773

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-330 <OHA>

A;Cross-references: EMBL:X63437

C;Superfamily: Mycobacterium avium alpha-antigen

Query Match

Best Local Similarity 87.0%; Score 1519; DB 2; Length 330;

Matches 274; Conservative 25; Mismatches 25; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
 Db 1 MTDLSEKVRWGRRLVGAATAATLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60

QY 61 DIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSLGSLVMPVGGQSSFFS 120
 Db 61 DIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSLGSLVMPVGGQSSFFS 120

QY 121 DWYSPACGKAGCQTYKWTETLTSELPAQLSANRAVKPTGSAALGLSMAGSSAMILAAVHP 180
 Db 121 DWYSPACGKAGCQTYKWTETLTSELPAQLSANRAVKPTGSAALGLSMAGSSAMILAAVHP 180

QY 181 QOPIYAGSLALLDPSQGMPSLIGLMDGAGGYKAADMWGPSSDPAWERNPTQOIPKL 240
 Db 181 QOPIYAGSLALLDPSQGMPSLIGLMDGAGGYKAADMWGPSSDPAWERNPTQOIPKL 240

QY 241 VANNTRLWVYCGNGCTNELGGANIPAEFLFNFRSSNLKFQDAYNAAGHNAVFNPNG 300
 DB 241 VGNTRLWVYCGNGCTSEVGGANWPAEFLFNFRSSNLKFQDAYNGAGGHNAVFNFNANG 300
 QY 301 THSWEYWGAGQLNAMKGDLOSSLGA 324
 DB 301 THSWEYWGAGQLNAMKPDQGLTGA 324
 RESULT 7
 G87162
 antigen 85A, mycolyltransferase [imported] - Mycobacterium leprae
 C/Species: Mycobacterium leprae
 C/Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
 C/Accession: G87162
 R/ Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho-
 R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
 eam, M.A.; Rutherford, K.M.
 Nature 409, 1007-1011, 2001
 A/Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq-
 A/Title: Massive gene decay in the leprosy bacillus.
 A/Reference number: A86909; MUID:21128732; PMID:11234002
 A/Accession: G87162
 A/Status: Preliminary
 A/Molecule type: DNA
 A/Residues: 1-327 <STC>
 A/Cross-references: GB:AL450380; NID:gl3093650; PIDN:CAC30983.1; GSPDB:GN00147
 C/Genetics:
 A/Gene: fbpB
 C/Superfamily: Mycobacterium avium alpha-antigen

Query Match	83.4%;	Score 1455;	DB 2;	Length 327;
Best Local Similarity	83.0%;	Pred. No. 4e-104;		
Matches 269;	Conservative 19;	Mismatches 34;	Indels 2;	Gaps 1;
QY	1	MTDYSKTRAWGRRLMIGTAAAVLPLGLVLGAGGAATAGAFSRPGLPVEYLQVPSMGR	60	
DB	1	MIDVSGKTRAWGRLLVVG--AAATLPSLISLAGGATASAFSRPGLPVEYLQVPSMGR	58	
QY	61	DIKQVQSQGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYYQSGLSIVMPVGQSSFSYS	120	
DB	59	SIKQVFQNGGNSPAVYLLDGLRAQDDYNGWDINTSAFEWYYQSGLSVMPVGQSSFSYS	118	
QY	121	DWYSPACKAGAGQCTYKWTFTLTSELPOWLSANRAVPTGCSRAAIGLSMAGSSAMILLAAYHP	180	
DB	119	DWYSPACKAGAGCTTYKWTFTLTSELPKWLSANRSVKSTGSAVWGLSMAGSSALLIAYHP	178	
QY	181	QOFIYAGSLSALLDPSQCGPSSLIGLAMDAGGYKAADWGPSSDPAWERNDPTQQIPKL	240	
DB	179	QOFIYAGSLSALLMDSQGLEPQLIGLAMDAGGYKAADWGPFPNDPAWRNDPILQAGKL	238	
QY	241	VANNTRLWYCGNGTSPNEIGGANIPAELEFNVRSSNLKFQDAYNACAGHNAVFNPNG	300	
DB	239	VANNTHLWYCGNGTSPSELGGTNVPAEFLNFVHGSNLKFQDAYNACAGHNAVFNIADG	298	
QY	301	THSWEYWGQALNAMKGDLOSSLGA	324	
DB	299	THSWEYWGQALNAMKPDQLQNTLMA	322	

RESULT 8
S34434
antigen 85-B - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 26-May-2000
C:Accession: S34434
R:de Mendonca Lima, L.; Content, J.; van Heuverswyn, H.; Degraeve, W.
Nucleic Acids Res. 19, 5789, 1991
A:Title: Nucleotide sequence of the gene coding for the 85-B antigen of Mycobacterium 1
A:Reference number: S34434; MUID:92051335; PMID:1945858
A:Accession: S34434
A>Status: preliminary

```
A:Molecule type: DNA
A:Residues: 1-327 <DEM>
A:Cross-references: EMBL:X60934; NID:G44380; PIDN:CAA43269.1; PID:G44381
C:Superfamily: Mycobacterium avium alpha-antigen

Query Match      83.3%; Score 1454; DB 2; Length 327;
Best Local Similarity 83.0%; Pred. No. 4.8e-104;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;
```

Qy	1	MTDVSKRIRAWGRRLMIGTAAAVVLCVLGGAGATAGAFSRPGLPVEYLQVPSPMGRR	60
Db	1	MIDVSGKIRAWGRWLLVG--AAATLSLISLAGATASAFSRPGLPVEYLQVFSEAMGR	58
		:::::	
Qy	61	DIKVQFOSGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYIOGSLIVMPVGQSFSYS	120
Db	59	TIKVQFONGGSPAVYLLDGLRAQDDYNGWDINTSAFEWYIOGSLVMVFGQSFSYS	118
		:::::	
Qy	121	DWTSPACGKAGCTYKWETELTSELPOWLKANRAVKPTGSAATGLSWAGSSAMTLAAHYH	180
Db	119	DWTSACGKAGCTTYKWETFLTSELPKWSANKRSVKTGSAVVGLSWAGSSALILAAHYH	178
		:::::	
Qy	181	QQFIYAGSLIALIDPSOGMGPSITGLAMGDAGGYKAADMWGSPSDPAWERNDPTQQPKL	240
Db	179	DQFIYAGSLALMDSSQGIEPQLIGLAMGDAGGYKAADMWGPPNDPAWRNDPTLIQAKL	238
		:::::	
Qy	241	VANNTRLVVYCGNGTPNELGGANIPAEFLNFVRESSNLKEODAYNAAGHNVAFFNPFG	300
Db	239	VANNTHLVVYCGNGTPSELGTNVPAEFLNFVHGSSLKPQDAYNCAGCHNAVFNLAADG	298
		:::::	
Qy	301	THSWEYMGAAQLNAMKGDLQSSLGA 324	
Db	299	THSWEYMGAAQLNAMKPLQNTLMA 322	

```

RESULT 9
S20038
fibronectin-binding antigen 85 precursor - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 19-Mar-1997 #sequence_revision 18-Jul-1997 #text_change 26-May-2000
C:Accession: S20038; S77953
R:Thole, J.E.R.; Schoeninger, R.; Janson, A.A.M.; Garbe, T.; Cornelisse, Y.E.; (C)
Mol. Microbiol. 6, 153-163, 1992
A:Title: Molecular and immunological analysis of a fibronectin-binding protein
A:Reference number: S20038; MUIB:92186705; PMID:1532043
A:Accession: S20038
A:Molecule type: DNA
A:Residues: 1-327 <THO>
A:Cross-References: EMBL:Z11666
R:Thole, J.E.R.
submitted to the EMBL Data Library, January 1992
A:Reference number: S77953
A:Accession: S77953
A:Molecule type: DNA
A:Residues: 1-63, 'F', 65-327 <THW>
A:Cross-References: EMBL:Z11666
C:Superfamily: Mycobacterium avium alpha-antigen
F1-38/Domain: signal sequence hstacus predicted <SIG>
F39-327/Product: fibronectin-binding antigen 85 #status predicted <MAT>
F55-266/Region: fibronectin-binding
F265-327/Region: fibronectin-binding

```

	Query Match	81.8%;	Score 1428;	DB 2;	Length 327;	
	Best Local Similarity	81.8%;	Pred. No. 4.7e-102;			
	Matches 265;	Conservative	20;	Mismatches 37;	Indels 2;	Gaps 1
QY	1	MTDVKIRAWGRRLMIGTAAAVVLPGLVLGGAGATAGAFSRPGLPVEYLQVPSPSMGR	60			
		:::::				
Dd	1	MTDVKIRAWGEWLLVG--AAATLPSLTSLAGGAATAFASRPGLPVEYLQVPSEAMGR	58			
QY	61	DIKVFQSGGNNSPAVILLDLGRAODDYGWDINTAFEFYYQSGLSIWMPVGGQSSEYS	120			
nB	59	STKVQI·ONGNGSPAIVLDGI·RAODDYGWDINTSAFEFYIQSGLSIWNPVGGQSSEYS	118			

QY 121 DWYSPACGKAGCTYKWTFTLSELPWLSANRAVKPTGSAAGLGLSMAGSAMILAAAYHP 180
 Db 119 DWYSPACGKAGCTYKWTFTLSELPKVCVSNRVSVKSTGRVVLGSLWAGSALILAAAYHP 178
 QY 181 QOFTYAGSLDPSQGMGSPSLIGLWAGDAGGYKAADMMGPPSDPAWERNNDPTQOIPKX 240
 Db 179 DQFTYAGSLDPSQGMGSPSLIGLWAGDAGGYKAADMMGPPNDPAWQNRNDPILQAGKL 238
 QY 241 VANNTRLVVYCGNGTNEIGGANIPAEFLNFVRSSNLKFDQDAYNAAGGNNAVFNPPNG 300
 Db 239 VANNTRLVVYCGNGTNEIGGANIPAEFLNFVRSSNLKFDQDAYNAAGGNNAVFNPPNG 300
 QY 301 THSWYWGAGLQNAKMGDQSSSLGA 324
 Db 299 THSWYWGAGLQNAKMGDQSSSLGA 322

RESULT 10

S10326
 alpha-antigen A, extracellular - Mycobacterium bovis
 C:Species: Mycobacterium bovis
 C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text_change 26-May-2000
 C:Accession: S10326; B41499
 R:De Wit, L.; de la Cuvelier, A.; Ooms, J.; Content, J.
 Nucleic Acids Res. 18, 3995, 1990
 A:Title: Nucleotide sequence of the 32 kDa-protein gene (antigen 85 A) of Mycobacterium
 A:Reference number: S10326; MUID:90326531; PMID:2197602
 A:Accession: S10326
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-338 <WIT>
 A:Cross-references: EMBL:X53034; NID:g44165; PIDN:CAA37206.1; PID:g44166
 R:Wiker, H.G.; Sletten, K.; Nagai, S.; Harboe, M.
 Infect. Immun. 58, 272-274, 1990
 A:Title: Evidence for three separate genes encoding the proteins of the mycobacterial an
 A:Reference number: A41499; MUID:90093478; PMID:2403534
 A:Accession: B41499
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 44-82 <WIK>
 C:Superfamily: Mycobacterium avium alpha-antigen

Query Match 80.6%; Score 1407; DB 2; Length 338;
 Best Local Similarity 78.9%; Pred. No. 2e-100;
 Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

QY 8 IRWGRRLMIGTAAAVLPLGLVLAGGAATAGAFSRPGLPVEYLQVPSMGREDIKVQFQ 67
 Db 11 VTGMSRRLVVGAVGAALVSLGAVGGTATAGAFSRPGLPVEYLQVPSMGREDIKVQFQ 70
 QY 68 SGNNSPAVYLLDGLRAODDYGNDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPAC 127
 Db 71 SGANSPLYLLDGLRAQDDFGWDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYQAPAC 130
 QY 128 GKAGCTYKWTFTLSELPWLSANRAVKPTGSAAGLGLSMAGSAMILAAAYHPQOIFYAG 187
 Db 131 GKAGCTYKWTFTLSELPWLVQANRHVKPTGSAVVLGSLMAASSALTALYHPQOIFYAG 190
 QY 188 SLGALLDPSQGMGSPSLIGLWAGDAGGYKAADMMGPPSDPAWERNNDPTQOIPKLVANNTR 247
 Db 191 AMSGLDPSQGMGPTLIGLWAGDAGGYKASDMWGPKEPAWQNRNDPILNVGKLIANNTRV 250
 QY 248 WYCGNGTNEIGGANIPAEFLNFVRSSNLKFDQDAYNAAGGNNAVFNPPNGTHSWEY 307
 Db 251 WYCGNGKPSDLGGNNLPKFLGFEVFTSNIKFDQDAYNAGGNGHGVDFPDGSDTHSWEY 310
 QY 308 GAQLNAKMGDQSSSLGA 324
 Db 311 GAQLNAKMGDQSSSLGA 327

RESULT 11

H70887

32K antigen fbpA precursor - Mycobacterium tuberculosis (strain H37RV)
 N:Alternate names: 30K native antigen; major protein antigen MPT45
 C:Species: Mycobacterium tuberculosis
 C:Date: 17-Jul-1998 #sequence revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: H70887; A37024; G60274; A54318
 R:Coale, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
 Nature 393, 537-544, 1998
 A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: H70887
 A:Status: nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-338 <COL>
 A:Cross-references: GB:AL022076; GB:AL123456; NID:g3256026; PIDN:CAA17868.1; PID:g295042
 R:Borreman, M.; De Wit, L.; Volckaert, G.; Ooms, J.; De Bruyn, J.; Huygen, K.; Van Voor
 Infect. Immun. 57, 3123-3130, 1989
 A:Title: Cloning, sequence determination, and expression of a 32-kilodalton-protein gene
 A:Reference number: A37024; MUID:89379378; PMID:2506131
 A:Accession: A37024
 A:Molecule type: DNA
 A:Residues: 1-23,25,'R',27-130,'R',132-291,'R',293-323,'HWVPRPTP',332-333,335-338 <BOR>
 A:Cross-references: GB:M27016; GB:X53898
 R:Nagai, S.; Wiker, H.G.; Harboe, M.; Kinomoto, M.
 Infect. Immun. 59, 372-382, 1991
 A:Title: Isolation and partial characterization of major protein antigens in the culture
 A:Reference number: A60274; MUID:91099989; PMID:1898899
 A:Accession: C60274
 A:Molecule type: protein
 A:Residues: 44-48 <NAG>
 R:Salata, R.A.; Sanson, A.J.; Malhotra, I.J.; Wiker, H.G.; Harboe, M.; Phillips, N.B.; D.
 J. Lab. Clin. Med. 118, 589-598, 1991
 A:Title: Purification and characterization of the 30,000 dalton native antigen of Mycobac
 tigen.
 A:Reference number: A54318; MUID:92078747; PMID:1720803
 A:Accession: A54318
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 44-63 <SAL>
 A:Experimental source: H37Ra
 A:Note: sequence extracted from NCBI backbone (NCBIP:68909)
 C:Genetics:
 A:Gene: fbpA
 C:Superfamily: Mycobacterium avium alpha-antigen
 F:1-43/Domain: signal sequence #status predicted <SIG>
 F:44-338/Product: 32K antigen fbpA #status experimental <MAT>

Query Match 80.6%; Score 1407; DB 2; Length 338;
 Best Local Similarity 78.9%; Pred. No. 2e-100;
 Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

QY 8 IRWGRRLMIGTAAAVLPLGLVLAGGAATAGAFSRPGLPVEYLQVPSMGREDIKVQFQ 67
 Db 11 VTGMSRRLVVGAVGAALVSLGAVGGTATAGAFSRPGLPVEYLQVPSMGREDIKVQFQ 70
 QY 68 SGNNSPAVYLLDGLRAODDYGNDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPAC 127
 Db 71 SGANSPLYLLDGLRAQDDFGWDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYQAPAC 130
 QY 128 GKAGCTYKWTFTLSELPWLSANRAVKPTGSAAGLGLSMAGSAMILAAAYHPQOIFYAG 187
 Db 131 GKAGCTYKWTFTLSELPWLVQANRHVKPTGSAVVLGSLMAASSALTALYHPQOIFYAG 190
 QY 188 SLGALLDPSQGMGSPSLIGLWAGDAGGYKAADMMGPPSDPAWERNNDPTQOIPKLVANNTR 247
 Db 191 AMSGLDPSQGMGPTLIGLWAGDAGGYKASDMWGPKEPAWQNRNDPILNVGKLIANNTRV 250
 QY 248 WYCGNGTNEIGGANIPAEFLNFVRSSNLKFDQDAYNAAGGNNAVFNPPNGTHSWEY 307
 Db 251 WYCGNGKPSDLGGNNLPKFLGFEVFTSNIKFDQDAYNAGGNGHGVDFPDGSDTHSWEY 310

[illegible]

RESULT 16
G70887
fbpC1 protein precursor - Mycobacterium tuberculosis
N:Alternate names: major antigen MPT51
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: G70887; A60274
R:Code, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: G70887

```

Query Match      28.6%; Score 498.5; DB 2; Length 301;
Best local Similarity 38.7%; Pred. No. 7e-31;
Matches 120; Conservative 39; Mismatches 124; Indels 27; Gaps 8;

15 LMIGTAAAVLPGVLGAGGAATAGASRPGLPVEYLQVPSPSMGRDITKVFQGGNNSP 74
16 LAVGVFAAAVL--LAGTAGAKAAG-----YESLNVPSNAMGRDIPVFAAGGPH-- 63

```

75 AVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSPACGKAGCQT 134
 116
 64 AVYLLDAFNAALDVSNWVTAGNMTTLGGRGISVAVAGGAYSMYTNWENDG-----S 116
 135 YKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAVHPQOFIYAGSLSLALLD 194
 117 KQWDTFUSSELPLWATKRGAPDCHAAVAGSOGGYAALAAALAHFDFRGFAGSLSGFVY 176
 195 PSQGMGPSLIGLAMGDAGGYKAADMWGPSSDPAMERNPTQOIPLKLVANNTRLVWYCGNG 254
 177 PSTNYNGAILAGLQQFGIDGNGMWGAPQLGRWKWHDPPYVHAILAQNTRVWVY---- 232
 255 TPNELGGANIPAEFLENFVRS--SNLKFQDAYNAGGNVAFNPPNGTHSWWYWGQOLN 312
 233 SPMTWGG-DIDA-MIGQAVASGMSRBEFYQYRSVGGHNGHDFSGGSDNGWGAAPQLA 290
 313 AMKGLQSSSL 322
 291 AMSGDIVGAI 300

RESULT 18
 S25184
 cspl protein - Corynebacterium glutamicum
 N:Alternate names: secreted protein PSI
 C:Species: Corynebacterium glutamicum
 C:Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Oct-1999
 C:Accession: S25184
 R:Joliff, G.; Mathieu, L.; Hahn, V.; Bayan, N.; Duchiron, F.; Renaud, M.; Shechter, E.;
 Mol. Microbiol. 6, 2349-2362, 1992
 A>Title: Cloning and nucleotide sequence of the cspl gene encoding PSI, one of the two
 cobacterium antigen 85 complex.
 A:Reference number: S25184; MUID:93023863; PMID:1406274
 A:Accession: S25184
 A:Molecule type: DNA
 A:Residues: 1-657 <JOL>
 A:Cross-references: EMBL:X66078; NID:g40486; PIDN:CAA46877.1; PID:g40487
 C:Genetics:
 A:Gene: cspl

Query Match 27.2%; Score 474; DB 2; Length 657;
 Best Local Similarity 36.1%; Pred. No. 1.3e-28;
 Matches 115; Conservative 40; Mismatches 140; Indels 24; Gaps 8;
 20 AAAYVLPGLVGLAGGAATAGAFRRPGLPVEYIQVPSPM-GRDIKVQ-----FQSGGN 71
 78 AGEVTPETAGLPDGRVVISAEWATSKHV-ILTIQSAAMPERPIKVQLLLPRDWYSSPNR 136
 72 NSPANYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSPACGKAG 131
 137 EFFEIWALDGRRAIERQSGWTIETNTEQYADKNALIVLPIGESSFYSDWEEPNNGK-- 194
 132 CQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAVHPQOFIYAGSLSLA 191
 195 --NYQWETFTQLAPILDKGFR-SNTDRAITGISMGGAIVNIAHTHPDMKFFVGSFSG 251
 192 LLDPSQGMGPSLIGLAMGDAGGYKAADMWGPSSDPAMERNPTQOIPLKLVANNTRLVWY 251
 252 YLDTTSAGMPIALSAALADAGGYDANAMWGPVGSERWQENDPKSNVDKL--KGTIYVSS 309
 252 GNGTPN-----ELGANIPAEFLENFVRSNLFQDAYNAGGNVAFNPPNGTHSW 304
 310 GNGADDFGKEDSVAGFANATGVGLEVISMTSCTTFVDRANQA-GVEVVASFRPSGVHSH 368
 305 EYWGAQINAMKGLQSSSLG 323
 369 EYQWFEMTQAPFHIANALG 387

RESULT 19
 S32111
 MPT51 protein - Mycobacterium leprae (fragment)
 C:Species: Mycobacterium leprae

C:Date: 06-Feb-1995 #sequence_revision 06-Feb-1995 #text_change 23-Mar-2001
 C:Accession: S32111
 R:Rinke de Wit, T.F.; Bekellie, S.; Osland, A.; Wieles, B.; Janson, A.A.M.; Thole, J.B.R.
 submitted to the EMBL Data Library, March 1993
 A:Description: The M.leprae antigen 85 complex gene family: identification of the genes
 A:Reference number: S32107
 A:Accession: S32111
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-220 <RIN>
 A:Cross-references: EMBL:Z21949; NID:g287923; PIDN:CAA79947.1; PID:g581345
 C:Genetics:
 A:Start codon: GTG
 C:Superfamily: Mycobacterium avium alpha-antigen

Query Match 21.1%; Score 367.5; DB 2; Length 220;
 Best Local Similarity 39.7%; Pred. No. 5.3e-21;
 Matches 87; Conservative 26; Mismatches 87; Indels 19; Gaps 4;
 15 LMIGTAAAVLPGLVGLAGGAATAGAFRRPGLPVEYIQVPSPMGRDIKVQFQSGGNNSP 74
 16 LAVGVFAAVL--LAGTAGNAKAG-----YESLMVPSNAMGRDIPVAFMAGGPH-- 63
 75 AVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSPACGKAGCQT 134
 64 AVYLLDAFNAALDVSNWVTAGNMTTLGGRGISVAVAGGAYSMYTNWENDG-----S 116
 135 YKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAVHPQOFIYAGSLSLALLD 194
 117 KQWDTFUSSELPLWATKRGAPDCHAAVAGSOGGYAALAAALAHFDFRGFAGSLSGFVY 176
 195 PSQGMGPSLIGLAMGDAGGYKAADMWGPSSDPAMERNPT 233
 177 PSTNYNGAILAGLQQFGIDGNGMWGAPQLGRWKWHD 215

RESULT 20
 B43603
 antigen 85-C - Mycobacterium bovis (strain BCG) (fragment)
 C:Species: Mycobacterium bovis
 C:Date: 12-Jan-1993 #sequence_revision 12-Jan-1993 #text_change 09-Jun-2000
 C:Accession: B43603
 R:Content, J.; de la Cuvelier, A.; De Wit, L.; Vincent-Levy-Frebault, V.; Ooms, J.;
 Infect. Immun. 59, 3205-3212, 1991
 A>Title: The genes coding for the antigen 85-C of Mycobacterium tuberculosis
 n of the gene coding for antigen 85-C of Mycobacterium tuberculosis.
 A:Reference number: A43603; MUID:9134869; PMID:1715324
 A:Accession: B43603
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-57 <CON>
 C:Superfamily: Mycobacterium avium alpha-antigen

Query Match 14.7%; Score 257; DB 2; Length 57;
 Best Local Similarity 88.0%; Pred. No. 3.3e-13;
 Matches 44; Conservative 4; Mismatches 12; Indels 0; Gaps 0;
 77 YLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSP 126
 1 YLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSP 50

RESULT 21
 T29115
 hypothetical protein - Streptomyces coelicolor
 C:Species: Streptomyces coelicolor
 C:Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
 C:Accession: T29115
 R:Parkhill, J.; Barrrell, B.G.; Rajandream, M.A.
 submitted to the EMBL Data Library, August 1998
 A:Reference number: Z17215
 A:Accession: T29115
 A:Status: preliminary; translated from GR/EMBL/DBJ

A;Molecule type: DNA
A;Residues: 1-342 <PAR>
A;Cross-references: EMB
C;Genetics:
A;Note: SC1F2.06

Query Match 13.4%; Score 233.5; DB 2; Length 342;
Best Local Similarity 28.2%; Pred. No. 1.6e-10;
Matches 100; Conservative 46; Mismatches 134; Indels 75; Gaps 21;

QY	15	LMIGTAAAVLP-GLVLIAGGAATAGAFSRPCLPVEY-----LQVPSMW-GRDI	62
Db	18	LVLAVLLAVLTPTAPEALAAKPBTCGA-TRHGAEVVAVTRVADRQVDLTVRSTALGRTV	76
QY	63	KVQF--QSGN-----NSPAYLLDGLRAQDYNGWDINTPAFEWYQSGLSIYMPV	112
Db	77	EVRLTTPDGWNPDDRHRHQHWPTLWLHGG--CCGDYTSWTSMTDVAETSLRDVLVWPE	134
QY	113	GGQSFYSYDWYSPACGAGCQTYKWETFLTSELQWL-----SANRAVKTGSAALGLS	166
Db	135	AGWNGWYSDWNH--GQGDPA--WETFTKELHLLERDWDGAGSNRW-----AGLS	183
QY	167	MAGSSAMILAAYHPQQF-----IYAGTSLALIDPSCMGPSLIGLWADAGY--KAADMW	220
Db	184	MGGQALLYAARHPGMPFRATAAFSGSAHPLNDES-----VRI MGFFAGQNDPLRW	237
QY	221	GPSDDPA-----WERNDPTQIQPKLVANNTRLWYCGNGT--PNELGAN--IPABF-LE	270
Db	238	G--DPVAQRGIWOAHDPFHLAKRL--RSIPVVLSCGDGTTGPDLPAGTASALEADFNQ	292
QY	271	NFVRSSNLKFDQAYNAGGHNAVFNEPPNCTHSEWYGAQINAMKGDLOSSLGAG	325
Db	293	NHALAAELK-----RVGARHVTHTFYCGTGHWAYRERELHSLPMLLCALRGV	341

RESULT 22

D70772
hypothetical protein Rv1288 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70772
R;Cole, S.-T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
C;Comor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70772
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-456 <COL>
A:Cross-references: GB:Z73419; GB:AL123456; NID:G3261573; PIDN:CAA97754.1; PID:e243272;
A:Experimental source: GB:Z73419; strain H37RV
C:Genetics:
C:Gene: Rv1288

Query Match 8.4%; Score 146.5; DB 2; Length 456;
Best Local Similarity 25.7%; Pred. No. 0.0011;
Matches 68; Conservative 34; Mismatches 86; Indels 77; Gaps 16;

[illegible]

```

QY 223 SSDPAWER-----NDPTQOIPIKLVAANTLWYICGNG-----TPNE-----LGGANIPA 266
Db 352 ---PLWDQARVSADNVERIDSY--RNKKIFLIVAGTSPDPANWFDSVNETVLGAGR--- 403
QY 267 EFLENFVRSSNLKP-QDAYNAAGGH 290
Db 404 EFRE---RLSNAGIPHESHEVPGGH 425

RESULT 23
E95071
tributylin esterase [imported] - Streptococcus pneumoniae (strain TIGR4)
C:Species: Streptococcus pneumoniae
C:Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 03-Aug-2001
C:Accession: E95071
R:Tettelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson
on, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Ho
nson, T.D.; Hickey, E.K.; Holt, I.E.
Science 293, 498-506, 2001
A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.;
A:Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneum
A:Reference number: A95000; MUID:21357209; PMID:114631916
A:Accession: E95071
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <KUR>
A:Cross-references: GB:AE005672; PIDN:AAK74766.1; PID:gl4972089; GSPDB:GN00164.1
A:Experimental source: strain TIGR4
C:Genetics:
A:Gene: SP0614

```

Query Match 8.2%; Score 143; DB 2; Length 259;
Best Local Similarity 27.0%; Pred. No. 0.001;
Matches 66; Conservative 28; Mismatches 104; Indels 46; Gaps 15;

```

QY 74 PAVYLLDGLRAQDDYNGNDINTPAFEWYQSLGVMPVGGQSFYSDWYSPACGKAGCQ 133
Db      :|||:      :|||:      :|||:      :|||:      :|||:
38 PVLVLLHGM--SGNHWSLKRFTNVERLLRGTNLIVWP-----NTSGWYTD--QYGF 88
QY 134 TYKWETILTSELPOWLS---ANRAVKPTGSAATGLSMAGSSAMILAAHYHPQFIYAGSL 190
Db      :|||:      :|||:      :|||:      :|||:      :|||:
89 YY---TALABELPQVLKRFPPNMTSKREKTFIAGLSMGYGCFKL-ALTINRFSHAASFS 144
QY 191 ALLDPSCMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERND-PTQIPIKLVANNTELWV 249
Db      :|||:      :|||:      :|||:      :|||:      :|||:
145 GALS-FQNFPSQNTL--GSPAYWRG--VFGEIRD--WTTSPYSLESLAKSKDKTKLWA 197
QY 250 YCGNGTENELGGANIPAEFL--ENFVRSSNLKQCDAYNAAGHNAVNFPPNGTHSWEYW 307
Db      :|||:      :|||:      :|||:      :|||:      :|||:
198 WCGE-----QDFLYEANNLAVKNLK-KLGFDTVYSHSA-----GTHEWYWW 237
QY 308 GAQL 311
Db      ||
238 EKQL 241

RESULT 24
C97939
tributylin esterase [imported] - Streptococcus pneumoniae (strain R6)
C:Species: Streptococcus pneumoniae
C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 22-Oct-2001
C:Accession: C97939
R:Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff,
e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhire,
Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskuna
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: C97939
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-259 <KUR>

```

A;Cross-references: GB:AE007317; PIDN:AAK99343.1; PID:g15458114; GSPDB:GN00174
C;Genetics:
A;Gene: estA

Query Match 8.2%; Score 143; DB 2; Length 259;
Best Local Similarity 27.0%; Pred. No. 0.001;
Matches 66; Conservative 28; Mismatches 104; Indels 46; Gaps 15;
QY 74 PAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCQ 133
Db 38 PVLVLLHGM--SGNHSWLKKTNVERLLRGTLNIVMP-----NTSNGWYDT--QYGF 88
QY 134 TYKWETFLTSELPWLS---ANRAVKPTGSAIGLSVAGSSAMILAAHYHPOQFIYAGSL 190
Db 89 YX---TALAELPOLVKRFFENMTSKREKTFIAGLSMGYGCFL-ALTNNRFSHAASF 144
QY 191 ALLDPSQMGPSLIGLAWDAGGYKAADMWGPSSDPAPERND-PTQOIPKLVANNRLIW 249
Db 145 GALS-FQNFSPESQNL--GSPAYWRG--VFGEIRD--WTTSPYSLSLAKSKDKTKLWA 197
QY 250 YCGNGTPTNELGGANIPAEFL--ENFVRSSNLKFQDAYNAAGHNAVFPPNGTHSWY 307
Db 198 WCGE-----QDFLYEANNLAVKNLK-KLGFQDVTYSHSA-----GTHWY 237
QY 308 GAQL 311
Db 238 BKQL 241

RESULT 25
AB1748
acetyltransferase homolog lin2527 [imported] - Listeria innocua (strain Clip11262)
C;Species: Listeria innocua
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 19-Apr-2002
C;Accession: AB1748
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AB1748
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-252 <GLA>
A;Cross-references: GB:AL592022; PIDN:CAC97754.1; PID:g16415049; GSPDB:GN00178
A;Experimental source: strain Clip11262
C;Genetics:
A;Gene: lin2527
C;Superfamily: conserved hypothetical protein YJL068c

Query Match 7.7%; Score 134; DB 2; Length 252;
Best Local Similarity 24.0%; Pred. No. 0.0048;
Matches 61; Conservative 40; Mismatches 85; Indels 68; Gaps 13;
QY 76 VYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCQ 135
Db 40 LYILHGL--SNNHTYVRNTNIERYATEKGLAVIMP--AADHSFYSNMVH---GRDFQ-- 91
QY 136 KWETFLTSELP-----QWLSANRAVKPTGSAIGLSVAGSSAMILAAHYHPOQFIYAGSL 191
Db 92 ----FVSTELPHVMKNFPLSKDKEDTFIA--GHSMMGYGAFKVALTFPEKFOAASMSG 145
QY 192 LLDPSQMGPSLIGLAWDAGGYKAADMWGPSSDPAPERNDPTQOIPKLVANNRLWY 251
Db 146 VMDINYL-----GGANIPAEF-----LENFVRSSNLKFQDAYNAAGHNAVFPP 174
QY 252 GNGTPTNEL-----GGANIPAEF-----LENFVRSSNLKFQDAYNAAGHNAVFPP 298
Db 175 --GTENDLPHLETLNLTNNVELPALFQNGCTEDFLYEDNLRFRD---FALAKNAPLEYRE 229

QY 299 N-GTHSWYWGCAQL 311
Db 230 GPGDHDWDFWDKSL 243

RESULT 26

AI1378
acetyltransferase homolog lmo2433 [imported] - Listeria monocytogenes (strain EGD-e)
C;Species: Listeria monocytogenes
C;Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 02-Aug-2002
C;Accession: AI1378
R;Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Science 294, 849-852, 2001
A;Authors: Kreft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Maok, C.; Schlueter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.; Title: Comparative genomics of Listeria species.
A;Reference number: AB1077; MUID:21537279; PMID:11679669
A;Accession: AI1378
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-252 <GLA>
A;Cross-references: GB:NC_003210; PIDN:CAD00511.1; PID:g16411921; GSPDB:GN00177
A;Experimental source: strain EGD-e
C;Genetics:
A;Gene: lmo2433
C;Superfamily: conserved hypothetical protein YJL068c

Query Match 7.6%; Score 133; DB 2; Length 252;
Best Local Similarity 23.8%; Pred. No. 0.0058;
Matches 61; Conservative 39; Mismatches 76; Indels 80; Gaps 13;
QY 76 VYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPACGK 129
Db 40 LYILHGL--SNNHTYVRNTNIERYATEKGLAVIMP--AADHSFYSNMVHGRDFE---- 91
QY 130 AGCQTYKWETFLTSELP-----QWLSANRAVKPTGSAIGLSVAGSSAMILAAHYHPOQFI 185
Db 92 ----FVSTELPHVMKNFPLSKDKEDTFIA--GHSMMGYGAFKVALTFPEKFOA 139
QY 186 AGSLALDPSQMGPSLIGLAWDAGGYKAADMWGPSSDPAPERNDPTQOIPKLVANN 245
Db 140 AASMSGVMDINYL-----KEDCFE--NFSTRATGEMASQT 174
QY 246 RLWVYCGNGTPTNEL-----GGANIPAEF-----LENFVRSSNLKFQDAYNAAGHNA 292
Db 175 ----GTENDLPHLETLNLTNNVELPALFQNGCTEDFLYEDNLRFRD---FALEXNA 223
QY 293 VFNFPFN--GTHSWY 307
Db 224 PLEYREGPGDHDWEP 239

RESULT 27

C86846
lipase [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C;Species: Lactococcus lactis subsp. lactis
C;Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001
C;Accession: C86846
R;Bolotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarre, K.; Weissenbach, J.; Ehrig, J.; Glaser, P.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, R.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.; Jones, L.M.; Karst, U.
Genome Res. 11, 731-753, 2001
A;Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis
A;Reference number: AB6625; MUID:21235186; PMID:11337471
A;Accession: C86846
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-258 <STO>
A;Cross-references: GB:AE005176; PID:g12724794; PIDN:AAK05869.1; GSPDB:GN00146
A;Experimental source: strain IL1403
C;Genetics:
A;Gene: yscE

```

Query Match          7.4%; Score 129.5; DB 2; Length 76;
Best Local Similarity 44.6%; Pred. No. 0.0027;
Matches 33; Conservative 7; Mismatches 21; Indels 13; Gaps 2;

      QY      25  LFLVLGLAGANTAGAFSPGLPVEVLQVPSPSMGRDIKVFQSGGNNSPAVYLLDGLRA 84
      Db      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
      15  LDNLVNVAAGST-----HRLMWPSRSMHRLIKVEFGGGPH--AWVLSDGIIA 61
      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
      QY      85  QDDYNGWDINTPAF 98
      Db      :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
      62  RDDYNGRDIHLPEVF 75

RESULT 29
C41499
alpha-antigen C, extracellular - Mycobacterium bovis (fragment)
N:Alternate names: antigen 85A; antigen P32
C:Species: Mycobacterium bovis
C>Date: 21-Apr-1992 #sequence_revision 21-Apr-1992 #text_change 09-Jun-2000
C:Accession: C41499; A61471
R:Wiker, H.G.; Sletten, K.; Nagai, S.; Harboe, M.
A:Title: Evidence for three separate genes encoding the proteins of the mycobacterial an
A:Reference number: A41499; MUID:90093478; PMID:2403534
A:Accession: C41499
A>Status: preliminary

```

Search completed: February 5, 2004, 17:40:45
Job time : 24.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:25:57 ; Search time 15.4762 Seconds
(without alignments)
987.561 Million cell updates/sec

Title: US-09-805-427A-2

Perfect score: 1745

Sequence: 1 MTDSVRKIRAWRRRLMIGTA.....YWGQQLNMGKDLQSLGAG 325

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127825

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1745	100.0	325	1	A85B MYCTU
2	1706	97.8	323	1	A85B MYCTU
3	1609	92.2	325	1	A85B MYCTU
4	1602	91.8	330	1	A85B MYCTU
5	1539	88.2	330	1	A85B MYCTU
6	1522	87.2	330	1	A85B MYCTU
7	1455	83.4	327	1	A85B MYCTU
8	1425	80.9	347	1	A85A MYCTU
9	1407	80.6	338	1	A85A MYCTU
10	1406	80.6	337	1	A85A MYCTU
11	1383.5	79.3	339	1	A85A MYCTU
12	1338.5	76.7	330	1	A85A MYCTU
13	1198	68.7	340	1	A85C MYCTU
14	1170	67.0	333	1	A85C MYCTU
15	1165	66.8	332	1	A85C MYCTU
16	625	35.8	139	1	A85A MYCTU
17	519	29.7	299	1	MP51 MYCTU
18	498.5	28.6	301	1	MP51 MYCTU
19	470	26.9	657	1	CSP1 CORGL
20	146.5	8.4	456	1	YC88 MYCTU
21	125	7.2	282	1	ESTD HUMAN
22	116	6.6	458	1	AGLE RHIME
23	109	6.2	827	1	XANP YANS2
24	106	6.1	414	1	PHAL PSELE
25	102	5.8	1106	1	GLI1 HUMAN
26	101	5.8	344	1	T082 HUMAN
27	100.5	5.8	426	1	Y967 CORGL
28	100	5.7	416	1	CBPB CANFA
29	99.5	5.7	555	1	WEPA EMENI
30	97.5	5.6	277	1	YALM BOOLI
31	97.5	5.6	2364	1	PGCA BOVIN
32	97	5.6	264	1	GUNS BRWCA
33	95.5	5.5	440	1	PORP_PSEAE

34	95.5	5.5	488	1	PHB ALCEFA
35	93.5	5.4	473	1	Y333 MYCLE
36	92.5	5.3	345	1	T082 MOUSE
37	91.5	5.2	299	1	Y038 YEAST
38	91.5	5.2	403	1	H136 ARATH
39	90.5	5.2	458	1	BPHA PSEPS
40	90.5	5.2	538	1	TCMA STRGA
41	90	5.2	563	1	IDS MOUSE
42	90	5.2	1101	1	GUNC CELFI
43	89	5.1	311	1	ISPE SYNEL
44	88.5	5.1	639	1	EY44 HUMAN
45	88	5.0	228	1	CE21 PEA
46	88	5.0	424	1	GUN1 RALSO
47	87	5.0	236	1	TENA BACSU
48	86.5	5.0	263	1	CB23 SOYEN
49	86	4.9	334	1	E13B HORVU
50	86	4.9	374	1	ADH3 KLULA
51	85.5	4.9	600	1	LAM2 HUMAN
52	85	4.9	299	1	YHXD BACSU
53	85	4.9	353	1	FAEB PENFN
54	84.5	4.8	268	1	CB28 PEA
55	84.5	4.8	339	1	DUSC MOUSE
56	84.5	4.8	444	1	RUMA XANCP
57	84.5	4.8	473	1	Y333 MYCTU
58	84.5	4.8	521	1	LAG3 MOUSE
59	84.5	4.8	901	1	A180 MOUSE
60	84.5	4.8	997	1	ATS7 HUMAN
61	84.5	4.8	1070	1	EMBC MYCLE
62	84	4.8	580	1	PTFB XANCP
63	84	4.8	725	1	GUNG CLOCE
64	84	4.8	864	1	ELS RAT
65	84	4.8	1123	1	V120 HSV11
66	83.5	4.8	511	1	GUNB_PSEFL
67	83.5	4.8	540	1	YH8M YEAST
68	83.5	4.8	1433	1	SUBF BACSU
69	83	4.8	269	1	CB22 PEA
70	83	4.8	361	1	PAX1 MOUSE
71	83	4.8	458	1	BPHA BURCE
72	83	4.8	474	1	SOXA HUMAN
73	83	4.8	700	1	PURL HALN1
74	83	4.8	1134	1	YML7 YEAST
75	82.5	4.7	500	1	AMPA CORGL
76	82.5	4.7	548	1	LAC1 PHLRA
77	82.5	4.7	722	1	GUNF CLOCE
78	82.5	4.7	1092	1	EMBA MYCSM
79	82	4.7	674	1	YL10 VIBCH
80	82	4.7	757	1	DHET GLDOX
81	82	4.7	860	1	ELS MOUSE
82	82	4.7	1365	1	KRES YEAST
83	82	4.7	3119	1	CAIC MOUSE
84	81.5	4.7	278	1	YEIG ECOLI
85	81.5	4.7	313	1	EBAG STRPL
86	81.5	4.7	345	1	ESTA STRSC
87	81.5	4.7	420	1	PEL BACSU
88	81.5	4.7	466	1	HYIN AGRRH
89	81.5	4.7	527	1	YF00 MYCPN
90	81.5	4.7	1654	1	OMPBP_RICRI

ALIGNMENTS

RESULT 1
A85B MYCTU
ID A85B MYCTU STANDARD; PRT: 325 AA.
AC P31952; Q9RM10;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-SEP-2003 (Rel. 4, Last annotation update)
DE Antigen 85-B precursor (85B) (Extracellular alpha-antigen) (Antigen 85 complex B) (A985B) (Mycoplasmal transferase 85B) (EC 2.3.1.-)
DE (Fibronectin-binding protein B) (30 kDa extracellular protein).
GN FBPB OR RV1886C OR MT1934 OR MTCY180.32.

OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN
 RP SEQUENCE FROM N.A.
 RC STRAIN=Erddmann;
 RX MEDLINE=95078461; PubMed=7987013;
 RA de Wit L., Palou M., Content J.;
 RT "Nucleotide sequence of the 85B-protein gene of Mycobacterium bovis
 BCG and Mycobacterium tuberculosis.";
 RL DNA Seq. 4:267-270(1994).
 [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Erddmann;
 RX MEDLINE=96333337; PubMed=8757831;
 RA Harth G., Lee B.Y., Wang J., Clemens D.L., Horwitz M.A.;
 RT "Novel insights into the genetics, biochemistry, and
 immunocytochemistry of the 30-kilodalton major extracellular protein
 of Mycobacterium tuberculosis.";
 RL Infect. Immun. 64:3038-3047(1996).
 [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eigmeier K., Gas S., Barry C.B. III, Tekaa F.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 complete genome sequence.";
 RL Nature 393:537-544(1998).
 [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 Bishai W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 [5]
 RN SEQUENCE OF 41-325 FROM N.A.
 RC STRAIN=ATCC 201 / H37Ra;
 RA Fan X.L., Xu Z.H.K., Bai G.C.H., Li Y.;
 RT "Cloning and construction of a eukaryotic expression vector containing
 Ag85B gene of Mycobacterium tuberculosis.";
 RL J. Mol. Cell. Immunol. 0:0-0(2000).
 CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
 THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
 MYCOLYTRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
 TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
 FOR MAINTAINING CELL WALL INTEGRITY.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
 N-TERMINAL OF C. GLUTAMICUM P51 PROTEIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.ebi.ac.uk/ebis/sib.ch/announce/>)
 or send an email to license@sib.ch.
 CC
 CC EMBL; X62398; CAA44269.1; -;
 CC EMBL; U38939; AAC44294.1; -;
 DR

DR EMBL; Z97193; CAB10044.1; -;
 DR EMBL; AEO07049; AAF46207.1; ALT_INT.
 DR EMBL; AF198032; AAF13448.1; -;
 DR PR; C70516; C70516
 DR PDB; 1FON; 28-MAR-01.
 DR PDB; 1FOP; 28-MAR-01.
 DR TIGR; MT1934; -;
 DR TubercuList; Rv1886c; -;
 DR InterPro; IPR000801; Esterase_put.
 DR Pfam; PF00756; Esterase; 1.
 KW Transferase; Acyltransferase; Antigen; Signal; Complete proteome;
 KW 3D-structure.
 FT SIGNAL 1 40 POTENTIAL.
 FT CHAIN 41 325 ANTIGEN 85-B.
 FT ACT_SITE 166 166 BY SIMILARITY.
 FT ACT_SITE 270 270 BY SIMILARITY.
 FT ACT_SITE 302 302 BY SIMILARITY.
 FT ACT_SITE 302 302 BY SIMILARITY.
 SQ SEQUENCE 325 AA; B993B5442FD5567D CRC64;
 Query Match 100.0%; Score 1745; DB 1; Length 325;
 Best Local Similarity 100.0%; Pred. No. 1.3e-124;
 Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTDSVKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFRRPGLPVEYLQVPSMGR 60
 DB 1 MTDSVKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFRRPGLPVEYLQVPSMGR 60
 QY 61 DIKQVQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSPYS 120
 DB 61 DIKQVQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSPYS 120
 QY 121 DWSPACGKAGCQTYKWTFTLTSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAHP 180
 DB 121 DWSPACGKAGCQTYKWTFTLTSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAHP 180
 QY 181 QQFIYAGSLALLDPSQGMPSLIGLAMDAGGYKAADWGPSSDPAMERNPTQIIPKL 240
 DB 181 QQFIYAGSLALLDPSQGMPSLIGLAMDAGGYKAADWGPSSDPAMERNPTQIIPKL 240
 QY 241 VANNTRLMVYCGNGTPNELGGANI PAEFLNFVRSNLKFDQAYNAAGSHNAVFPPNG 300
 DB 241 VANNTRLMVYCGNGTPNELGGANI PAEFLNFVRSNLKFDQAYNAAGSHNAVFPPNG 300
 QY 301 THSWEYGAQLNAMKGDLSLQSLGAG 325
 DB 301 THSWEYGAQLNAMKGDLSLQSLGAG 325
 RESULT 2
 A85B MYCBO STANDARD; PRT; 323 AA.
 ID A35E MYCBO
 AC P12542;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antigen 85-B precursor (85B) (Extracellular alpha-antigen) (Antigen 85
 complex B) (Ag85B) (Mycobyl transferase 85B) (EC 2.3.1.-)
 DE (Fibronectin-binding protein B).
 GN FBPP.
 OS Mycobacterium bovis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1765;
 [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 41-70.
 RC STRAIN=BCG;
 RX MEDLINE=88314872; PubMed=2842287;
 RA Matsuo K., Yamaguchi R., Yamazaki A., Tasaka H., Yamada T.;
 RT "Cloning and expression of the Mycobacterium bovis BCG gene for
 extracellular alpha antigen.";
 RL J. Bacteriol. 170:3847-3854(1988).
 [2]
 RN SEQUENCE FROM N.A.

```

RC STRAIN=BCG / Paris 1173 P2;
RX MEDLINE=95078461; PubMed=7987013;
RA de Wit L., Palcou M., Content J.;
RT "Nucleotide sequence of the 85B-protein gene of Mycobacterium bovis
RL BCG and Mycobacterium tuberculosis.";
RL DNA Seq. 4,267-270(1994).
CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
CC MYCOLYTRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
CC N-TERMINAL OF C.GLUTAMICUM PS1 PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M21839; AAA25359.1; -.
CC DR EMBL; X62397; CAA44268.1; -.
CC DR PIR; A32348; A32348.
CC DR PIR; S29663; S29663.
CC DR HSSP; P31953; 1DQY.
CC DR InterPro; IPR000801; Esterase_put.
CC DR Pfam; PF00756; Esterase; 1.
CC KW Transferase; Acyltransferase; Antigen; Signal.
CC FT SIGNAL 1 40
CC FT CHAIN 41 323 ANTIGEN 85-B.
CC FT ACT_SITE 166 166 BY SIMILARITY.
CC FT ACT_SITE 268 268 BY SIMILARITY.
CC FT ACT_SITE 300 300 BY SIMILARITY.
CC FT ACT_SITE 302 302 BY SIMILARITY.
CC FT CONFLICT 200 200 G -> GPS (IN REF. 2).
CC FT CONFLICT 283 284 KP -> NA (IN REF. 2).
CC FT SEQUENCE 323 AA; 34403 MW; 26C2A1BFID0464BA CRC64;
CC -----
Query Match 97.8%; Score 1706; DB 1; Length 323;
Best Local Similarity 98.5%; Pred. No. 1.2e-121;
Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
QY 1 MTDVSKIRAWGRRLMIGTAAAVLPGLVGLAGGATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MTDVSKIRAWGRRLMIGTAAAVLPGLVGLAGGATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
Db 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
QY 121 DWYSPACGKAGCQTYKWTETLTSELPAWLSANRAVKPTGSAALGLSMAGSSAMILAAYHP 180
Db 121 DWYSPACGKAGCQTYKWTETLTSELPAWLSANRAVKPTGSAALGLSMAGSSAMILAAYHP 180
QY 181 QQFIYAGLSALLDPSQGMGSLIGLAMDAGGYKAADMWGPSSDPANERNPTQOIPKL 240
Db 181 QQFIYAGLSALLDPSQGMG--LIGLAMDAGGYKAADMWGPSSDPANERNPTQOIPKL 238
QY 241 VANNTRLWYCGNGTNPENELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPFG 300
Db 239 VANNTRLWYCGNGTNPENELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPFG 298
QY 301 THSWEYGAQLNAMKGDIOSSLAG 325
Db 299 THSWEYGAQLNAMKGDIOSSLAG 323

```

RESULT 3

A85B MYCA

ID A85B MYCA

AC P21160;

```

DT 01-MAY-1991 (Rel. 18, Created)
DT 01-MAY-1991 (Rel. 18, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antigen 85-B precursor (85B) (Extracellular alpha-antigen) (Antigen 85
DE complex B) (Ag85B) (Mycolyl transferase 85B) (EC 2.3.1.-)
DE (Fibronectin-binding protein B).
DE FBPB.
GN Mycobacterium kansasii.
OS Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1768;
RV [1]
RX SEQUENCE FROM N.A.
RX MEDLINE=90129315; PubMed=2404875;
RA Matsuo K., Yamaguchi R., Yamazaki A., Tasaka H., Terasaka K.,
RA Yamada T.;
RT "Cloning and expression of the gene for the cross-reactive alpha
RT antigen of Mycobacterium kansasii.";
RL Infect. Immun. 58:550-556(1990).
CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
CC MYCOLYTRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
CC N-TERMINAL OF C.GLUTAMICUM PS1 PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X53897; CAA37868.1; -.
CC DR PIR; A37185; A37185.
CC DR HSSP; P31953; 1DQY.
CC DR InterPro; IPR000801; Esterase_put.
CC DR Pfam; PF00756; Esterase; 1.
CC KW Transferase; Acyltransferase; Antigen; Signal.
CC FT SIGNAL 1 40
CC FT CHAIN 41 325 ANTIGEN 85-B.
CC FT ACT_SITE 166 166 BY SIMILARITY.
CC FT ACT_SITE 270 270 BY SIMILARITY.
CC FT ACT_SITE 302 302 BY SIMILARITY.
CC FT SEQUENCE 325 AA; 34323 MW; 5F2281BCC48AE30D CRC64;
CC -----
Query Match 92.2%; Score 1609; DB 1; Length 325;
Best Local Similarity 89.8%; Pred. No. 2.4e-114;
Matches 291; Conservative 20; Mismatches 13; Indels 0; Gaps 0;
QY 1 MTDVSKIRAWGRRLMIGTAAAVLPGLVGLAGGATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MTDVSKIRAWGRRLMIGTAAAVLPGLVGLAGGATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
Db 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
QY 121 DWYSPACGKAGCQTYKWTETLTSELPAWLSANRAVKPTGSAALGLSMAGSSAMILAAYHP 180
Db 121 DWYSPACGKAGCQTYKWTETLTSELPAWLSANRAVKPTGSAALGLSMAGSSAMILAAYHP 180
QY 181 QQFIYAGLSALLDPSQGMGSLIGLAMDAGGYKAADMWGPSSDPANERNPTQOIPKL 240
Db 181 QQFIYAGLSALLDPSQGMGSLIGLAMDAGGYKAADMWGPSSDPANERNPTQOIPKL 240
QY 241 VANNTRLWYCGNGTNPENELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPFG 300
Db 241 VANNTRLWYCGNGTNPENELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPFG 300

```

181 QQFIYAGSLALDPQSGMGPSLIGLAMDAGYKADMGWSSDPARWENDTQOIPKL 240
 181 QQFIYAGSLALMDPSQGMGPSLIGLAMDAGYKASDMWSSDPARWENDTQOIPKL 240
 241 VANNRLWYCGNGTNEELGGANIPEEFLENFRSSNLKFQDAYNAGGHNAVFNPFG 300
 241 VGNTRLWYCGNGTPELGGANPAEFLENFRSSNLKFQDAYNAGGHNAVFHPDNG 300
 301 THSWEYGAQNLNMGDLQSSLA 324
 301 THSWEYGAQNLNMGDLQTLGA 324

RESULT 5
 A85B_MYCIT STANDARD; PRT; 330 AA.
 ID A85B_MYCIT
 AC Q49575; P94938;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Antigen 85-B precursor (85B) (Extracellular alpha-antigen) (Antigen 85
 DE complex B) (Ag85B) (Mycobyl transferase 85B) (EC 2.3.1.-)
 DE (Fibronectin-binding protein B).
 GN FPBP.
 OS Mycobacterium intracellulare.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1767;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13950;
 RX MEDLINE=94071912; PubMed=8250904;
 RA Kitaura H., Ohara N., Matsuo T., Tasaka H., Kobayashi K., Yamada T.;
 RT "Cloning, sequencing and expression of the gene for alpha antigen
 RT from Mycobacterium intracellulare and use of PCR for the rapid
 RT identification of Mycobacterium intracellulare";
 RL Biochem. Biophys. Res. Commun. 196:1466-1473(1993).
 CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
 CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
 CC MYCOLYTRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
 CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
 CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
 CC N-TERMINAL OF C. GLUTAMICUM P51 PROTEIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D16546; BAA03981.1; -;
 CC EMBL; D14253; BAA03243.1; -;
 CC PIR; JN0897; JN0897.
 CC HSP; P31953; 1DQY.
 CC InterPro; IPR000801; Esterase_put.
 CC Pfam; PF00756; Esterase; 1.
 CC Transferrase; Acyltransferase; Antigen; Signal.
 CC SIGNAL 1 40 BY SIMILARITY.
 CC CHAIN 41 330 ANTIGEN 85-B.
 CC ACT_SITE 166 166 BY SIMILARITY.
 CC ACT_SITE 270 270 BY SIMILARITY.
 CC ACT_SITE 302 302 BY SIMILARITY.
 CC ACT_SITE 302 302 BY SIMILARITY.
 CC CONFLICT 47 48 PV -> EF (IN REF. 1; BAA03243).
 CC CONFLICT 310 310 Q -> S (IN REF. 1; BAA03243).
 CC SEQUENCE 330 AA; 34507 MW; 83090670FABAE0D5 CRC64;

Query Match 88.2%; Score 1539; DB 1; Length 330;
 Best Local Similarity 85.8%; Pred. No. 4.6e-109;
 Matches 278; Conservative 23; Mismatches 23; Indels 0; Gaps 0;

301 THSWEYGAQNLNMGDLQSSLA 324
 301 THSWEYGAQNLNMGDLQSLGA 324

RESULT 4
 A85B_MYCSC STANDARD; PRT; 330 AA.
 ID A85B_MYCSC
 AC Q50397;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antigen 85-B precursor (85B) (Extracellular alpha-antigen) (Antigen 85
 DE complex B) (Ag85B) (Mycobyl transferase 85B) (EC 2.3.1.-)
 DE (Fibronectin-binding protein B).
 GN FPBP.
 OS Mycobacterium scrofulaceum.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1783;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 19981;
 RX MEDLINE=94323707; PubMed=8047837;
 RA Takano M., Ohara N., Mizuno A., Yamada T.;
 RT "Cloning, sequencing and expression in Escherichia coli of the gene
 RT for alpha antigen from Mycobacterium scrofulaceum";
 RL Scand. J. Immunol. 40:165-170(1994).
 CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
 CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
 CC MYCOLYTRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
 CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
 CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
 CC N-TERMINAL OF C. GLUTAMICUM P51 PROTEIN.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; D26187; BAA05332.1; -;
 CC HSP; P31953; 1DQY.
 CC InterPro; IPR000801; Esterase_put.
 CC Pfam; PF00756; Esterase; 1.
 CC Transferrase; Acyltransferase; Antigen; Signal.
 CC SIGNAL 1 40 BY SIMILARITY.
 CC CHAIN 41 330 ANTIGEN 85-B.
 CC ACT_SITE 166 166 BY SIMILARITY.
 CC ACT_SITE 270 270 BY SIMILARITY.
 CC ACT_SITE 302 302 BY SIMILARITY.
 CC SEQUENCE 330 AA; 34943 MW; 09D6E04B5E9ED221 CRC64;

Query Match 91.8%; Score 1602; DB 1; Length 330;
 Best Local Similarity 89.5%; Pred. No. 8.2e-114;
 Matches 290; Conservative 22; Mismatches 12; Indels 0; Gaps 0;

1 MTDSVKIRAWGRRLMTGTAANVVLPGVLGAGNATAGAFSRPLPVEYLQVPSMGR 60
 1 MTDLSSKVRAGRRLVGTAAVTLPLGLLGLAGAPTAGAFSRPLPVEYLQVPSMGR 60
 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFWYQSLGSLVMPVQSSSFYS 120
 61 NIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFWYQSLGSLVMPVQSSSFYS 120
 121 DWYSPACGKAGCTYKWTFTLSELTPQLWLSANRAVKPTGSAAGLSWAGSSAMTILAYHP 180
 121 DWYSPACGKAGCTYKWTFTLSELTPQLWLSANRAVKPTGSAAGLSWAGSSAMTILAYHP 180

CC MYCOLYTRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
CC
CC -1- SUBCELLULAR LOCATION: Secreted.
CC
CC -1- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
CC N-TERMINAL OF C. GLUTAMICUM P51 PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; X60934; CAA43269.1; --
CC EMBL; AL049571; CAB40285.1; --
CC EMBL; AL583924; CAC30983.1; --
CC F01; G87162; G87162.
CC F01; S34434; S34434.
CC HSP; P31953; IDQY.
CC Leproma; ML2028; --
CC InterPro; IPR000801; Esterase; 1.
CC Pfam; PF00756; Esterase; 1.
CC Transferase; Acyltransferase; Antigen; Signal; Complete proteome.
CC
CC SIGNAL 1 38 POTENTIAL.
CC CHAIN 39 327 ANTIGEN 85-B.
CC ACT SITE 164 164 BY SIMILARITY.
CC ACT SITE 268 268 BY SIMILARITY.
CC ACT SITE 300 300 BY SIMILARITY.
CC CONFLICT 59 59 S -> T (IN REF. 1).
CC SEQUENCE 327 AA; 34807 MW; 2B0EB9B37A8F985C CRC64;
CC
CC
CC Query Match 83.4%; Score 1455; DB 1; Length 327;
CC Best Local Similarity 83.0%; Pred. No. 9.9e-103;
CC Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;
CC
CC QY 1 MTDVSRKTRAGRRIMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
CC DB 1 MIDVSGKIRAWGRVLVG--AAATPLSLISLAGGAATASAFSRPGLPVEYLQVPSMGR 58
CC
CC QY 61 DIKVOFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFY 120
CC DB 59 SIKVOFQGGNGSPAVYLLDGLRAQDDYNGWDINTSAFEWYQSGLSIVMPVGGQSSFY 118
CC
CC QY 121 DWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
CC DB 119 DWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 178
CC
CC QY 181 QDFIYAGSLSDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPAMERNPTQIQPKI 240
CC DB 179 QDFIYAGSLSDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPAMERNPTQIQPKI 238
CC
CC QY 241 VANNTRFWYCGNGTNPNEGGLNIPAEFLNFVRSSNLKFDAYNAAGGHNVAFFPPNG 300
CC DB 239 VANNTRFWYCGNGTNPNEGGLNIPAEFLNFVRSSNLKFDAYNAAGGHNVAFFPPNG 298
CC
CC QY 301 THSWYWGQAQLNMGKDLQSSSLGA 324
CC DB 299 THSWYWGQAQLNMGKDLQNTLMA 322
CC
CC RESULT 8
CC A85A MYCAV STANDARD; PRT; 347 AA.
CC ID A85A MYCAV
CC AC O52956;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 16-OCT-2001 (Rel. 40, Last annotation update)
CC DE Antigen 85-A precursor (85A) (Antigen 85 complex A) (Ag85A) (Mycoly
CC transferase 85A) (EC 2.3.1.-) (Fibronectin-binding protein A).
CC GN BPBA.
CC OS Mycobacterium avium.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1764;
RN
RP SEQUENCE FROM N.A.
RX MEDLINE=97427949; PubMed=9284137;
RA Ohara N., Ohara-Wada N., Kitaura H., Nishiyama T., Matsumoto S.,
RA Yamada T.;
RA "Analysis of the genes encoding the antigen 85 complex and MPT51 from
RT Mycobacterium avium";
RL Infect. Immun. 65:3680-3685 (1997).
CC -1- FUNCTION: PROTEIN OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
CC MYCOLYTRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
CC
CC -1- SUBCELLULAR LOCATION: Secreted
CC -1- PFM: CONTAINS ONE DISULFIDE BRIDGE (BY SIMILARITY).
CC -1- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
CC N-TERMINAL OF C. GLUTAMICUM P51 PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D78144; BAA24156.1; --
CC HSP; P31953; IDQY.
CC InterPro; IPR000801; Esterase; 1.
CC Pfam; PF00756; Esterase; 1.
CC Transferase; Acyltransferase; Signal; Antigen.
CC
CC SIGNAL 1 43 POTENTIAL.
CC CHAIN 44 347 ANTIGEN 85-A.
CC ACT SITE 169 169 BY SIMILARITY.
CC ACT SITE 273 273 BY SIMILARITY.
CC ACT SITE 305 305 BY SIMILARITY.
CC SEQUENCE 347 AA; 36095 MW; DFAFF3DE504C208E CRC64;
CC
CC Query Match 80.9%; Score 1412.5; DB 1; Length 347;
CC Best Local Similarity 78.9%; Pred. No. 1.7e-99;
CC Matches 258; Conservative 26; Mismatches 40; Indels 3; Gaps 1;
CC
CC QY 1 MTDVSR---KIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPS 57
CC DB 1 MTLVDRLRGAVAGMPRRLLVGVGAAGALLSGLIGAVGSSATAGAFSRPGLPVEYLQVPSAA 60
CC
CC QY 58 MGRDVKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSS 117
CC DB 61 MGRDVKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSS 120
CC
CC QY 118 FYSDWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAA 177
CC DB 121 FYSDWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAA 180
CC
CC QY 178 YHPQDFIYAGSLSDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPAMERNPTQIQ 237
CC DB 181 YHPQDFIYAGSLSDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPAMERNPTQIQ 240
CC
CC QY 238 PKLVANNTRFWYCGNGTNPNEGGLNIPAEFLNFVRSSNLKFDAYNAAGGHNVAFFPP 297
CC DB 241 PKLVANNTRFWYCGNGTNPNEGGLNIPAEFLNFVRSSNLKFDAYNAAGGHNVAFFPP 300
CC
CC QY 298 PNGTHSWYWGQAQLNMGKDLQSSSLGA 324
CC DB 301 ANGTHDWPYWGQAQLNMGKDLQSSSLGA 327
CC
CC RESULT 9
CC A85A MYCTU STANDARD; PRT; 338 AA.
CC ID A85A MYCTU


```

DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antigen 85-A precursor (85A) (Antigen 85 complex A) (Ag85A) (Mycoly1
DE transferase 85A) (EC 2.3.1.-) (Fibronectin-binding protein A).
GN FBPA.
OS Mycobacterium ulcerans.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1909;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ITM 5150;
RA Tanghe A.J.;
RT "Protective efficacy of DNA vaccine encoding antigen 85A from M. bovis
RT BCG against Buruli Ulcer.";
RL Thesis (2001), Universite Libre de Bruxelles, Belgium.
CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
CC MYCOLYTRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted (by similarity).
CC -!- PTM: CONTAINS ONE DISULFIDE BRIDGE (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
CC N-TERMINAL OF C. GLUTAMICUM PSI PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ300576; CAC40861.1; -.
DR InterPro; IPR000801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
KW Transferase; Acyltransferase; Signal; Antigen.
FT SIGNAL 1 42 BY SIMILARITY.
FT CHAIN 43 337 ANTIGEN 85-A.
FT ACT_SITE 169 169 BY SIMILARITY.
FT ACT_SITE 273 273 BY SIMILARITY.
FT ACT_SITE 305 305 BY SIMILARITY.
SQ SEQUENCE 337 AA; 35643 MW; 99022B3407E5773 CRC64;

Query Match 80.6%; Score 1406; DB 1; Length 337;
Best Local Similarity 80.4%; Pred. No. 5e-99;
Matches 251; Conservative 31; Mismatches 30; Indels 0; Gaps 0;

QY 13 RRLMIGTAAAVLPGLVGLAGGAATAGAFSPRLPVEYLQVPSMGRDIKVFQSGGNN 72
DB 16 RRLMVGAVGAALLSLGLVGVGGATASAFSPRLPVEYLQVPSMGRNIKVFQSGGAN 75

QY 133 QTYKWTFTLSELQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHPDQFVYAGLSAL 192
DB 136 TTYKWTFTLSELQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHPDQFVYAGLSAL 195

QY 193 LDPSQMGPSLIGLAMGDAGGYKAADWGPSPDPAWRNDPTQIQPKLVANNTLWYCG 252
DB 196 LDPSQMGPSLIGLAMGDAGGYKAADWGPSPDPAWRNDPTQIQPKLVANNTLWYCG 254
DB 316 AMRPDLQTLGA 327

```

```

RESULT 11
A85A_MYCGO STANDARD; PRT; 339 AA.
AC O06052;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antigen 85-A precursor (85A) (Antigen 85 complex A) (Ag85A) (Mycoly1
DE transferase 85A) (EC 2.3.1.-) (Fibronectin-binding protein A).
GN FBPA.
OS Mycobacterium gordonae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1778;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14470;
RA Dumonceaux M.; Fauville-Dufaux M.; Ooms J.; de Wit L.; Content J.;
RT Cloning and sequencing of the 85A antigen from Mycobacterium
RT gordonae and its use for the specific PCR identification of these
RT mycobacteria.
RT Submitted (JAN-1997) to the EMBL/GenBank/DBJ databases.
RL -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
RL THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
RL MYCOLYTRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
RL TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
RL FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
RL -!- SUBCELLULAR LOCATION: Secreted.
RL -!- PTM: CONTAINS ONE DISULFIDE BRIDGE (BY SIMILARITY).
RL -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
RL N-TERMINAL OF C. GLUTAMICUM PSI PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Y10378; CAA71406.1; -.
DR HSP; P31953; IDQY.
DR InterPro; IPR000801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
KW Transferase; Acyltransferase; Signal; Antigen.
FT SIGNAL 1 43 POTENTIAL.
FT CHAIN 44 339 ANTIGEN 85-A.
FT ACT_SITE 169 169 BY SIMILARITY.
FT ACT_SITE 272 272 BY SIMILARITY.
FT ACT_SITE 304 304 BY SIMILARITY.
SQ SEQUENCE 339 AA; 35474 MW; 909B7E0E3969E2DE CRC64;

Query Match 79.3%; Score 1383.5; DB 1; Length 339;
Best Local Similarity 79.5%; Pred. No. 2.5e-97;
Matches 248; Conservative 33; Mismatches 30; Indels 1; Gaps 1;

QY 13 RRLMIGTAAAVLPGLVGLAGGAATAGAFSPRLPVEYLQVPSMGRDIKVFQSGGNN 72
DB 16 RRLMVGAVGAALLSLGLVGVGGATASAFSPRLPVEYLQVPSMGRNIKVFQSGGAN 75

QY 73 SPAYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSPACGKAGC 132
DB 76 SPAYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSPACGKAGC 135

QY 133 QTYKWTFTLSELQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHPDQFVYAGLSAL 192
DB 136 TTYKWTFTLSELQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHPDQFVYAGLSAL 195

QY 193 LDPSQMGPSLIGLAMGDAGGYKAADWGPSPDPAWRNDPTQIQPKLVANNTLWYCG 252
DB 196 LDPSQMGPSLIGLAMGDAGGYKAADWGPSPDPAWRNDPTQIQPKLVANNTLWYCG 254
DB 253 NGTPELGGANI PAEFLNFRSRLKFDQAYNAAGHNAVFPPNGTHSWYWGQAQLN 312

```

Db 255 DGKPSDLGGNNLPKFLGFRVTSNLKFOEAYNGAGGHNVAFFDANGTHDFWYGAAPVQ 314
 QY 313 AMKGDLOSSLGA 324
 Db 315 AMKGDLOSSLGA 326

RESULT 12
 ID A85A MYCLE STANDARD; PRT; 330 AA.
 AC Q05861;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antigen 85-A precursor (85A) (Antigen 85 complex A) (Ag85A) (Mycoly1
 GN FBPA OR ML0097; (Fibronectin-binding protein A).
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Thai53;
 RX MEDLINE=95130993; PubMed=7829901;
 RA Yin Y.;
 RT "Molecular cloning of alpha antigen like protein gene of
 RT Mycobacterium leprae and its over production in Escherichia coli.";
 RL Kansenshogaku Zasshi 68:1330-1337(1994).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA de Mendonca-Lima L.;
 RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrall B.G.;
 RT "Massive gene decay in the leprosy bacillus";
 RL Nature 409:1007-1011(2001).
 RN [4]
 RP SEQUENCE OF 143-330 FROM N.A.
 RX MEDLINE=93366419; PubMed=8359887;
 RA Rinke de Wit T.F., Bekelie S., Osland A., Wiele B.,
 RA Janson A.A.M., Thole J.E.R.;
 RT "The Mycobacterium leprae antigen 85 complex gene family:
 RT identification of the genes for the 85A, 85C, and related MP751
 RT proteins.";
 RL Infect. Immun. 61:3642-3647(1993).
 CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
 CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
 CC MYCOLITRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
 CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
 CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- PTM: CONTAINS ONE DISULFIDE BRIDGE (BY SIMILARITY).
 CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
 CC N-TERMINAL OF C. GLUTAMICUM P51 PROTEIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D43841; BAA07864.1; -;
 CC EMBL; M90648; AAA91864.1; -;
 CC EMBL; AL583917; CAC29605.1; -;
 CC EMBL; Z21950; CAA79948.1; -;
 CC DR PIR; A86921; A86921.
 CC DR PIR; S32107; S32107.
 CC DR HSSP; P31953; LDQY.
 CC DR Leproma; ML0097; -;
 CC DR InterPro; IPR000801; Esterase_put.
 CC DR Pfam; PF00756; Esterase; 1.
 CC KW Transferase; Acyltransferase; Signal; Antigen; Complete proteome.
 CC FT SIGNAL 1 42 POTENTIAL.
 CC FT CHAIN 43 330 ANTIGEN 85-A.
 CC FT ACT_SITE 168 168 BY SIMILARITY.
 CC FT ACT_SITE 272 272 BY SIMILARITY.
 CC FT ACT_SITE 304 304 BY SIMILARITY.
 CC FT CONFLICT 149 149 Q -> E (IN REF. 1).
 CC SQ SEQUENCE 330 AA; 35411 MW; 170C7C98C98FC5EC CRC64;

Query Match 76.7%; Score 1338.5; DB 1; Length 330;
 Best Local Similarity 76.9%; Pred. No. 6.1e-94;
 Matches 240; Conservative 30; Mismatches 41; Indels 1; Gaps 1;
 QY 13 RRLMIGTAAAVLPCLVGLAGGAATAGAFSPGLPVEYLQVPSMGRDIIKVFQSGNN 72
 Db 16 RRLVVEAMGVALLSALIGV-GSAPAEAFSPGLPVEYLQVPSMGRDIIKVFQSGNN 74
 QY 73 SPAVYLLDGLRAQDDYNGWDINTPAFWYQSGLSIVMPVGQSGSFYSDWYSPAGKAGC 132
 Db 75 SPALYLLDGLRAQDDYNGWDINTPAFWYQSGLSIVMPVGQSGSFYSDWYSPAGKAGC 134
 QY 133 QTYKWETFLTSELDPQLSANRAVKPTGSAALGLSMAGSSAMILAAHYHPOQFIYAGSL 192
 Db 135 QTYKWETFLTSELDPQLSANRAVKPTGSAALGLSMAGSLALTAIYHDPQFIYVGS 194
 QY 193 LDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPWERNPTQOIPKLVANNTLWYVCG 252
 Db 195 LDPSNMGPSLIGLAMGDAGGYKAADMWGPSTDPWKRNDPTVNVGTLIANNTRIWYVCG 254
 QY 253 NGTPELGGANIPAEFLNFVRSSNLKFDAYNAAGHNVAFFPPNGTHSWYWGQLN 312
 Db 255 NGKPTELGGNNLPAKLEGLVRSNIKFDQYAGGHNVAFFPPDSGTHSWYWGQLN 314
 QY 313 AMKGDLOSSLGA 324
 Db 315 DMKPDLOQYLGA 326

RESULT 13

A85C MYCTU

ID A85C MYCTU STANDARD; PRT; 340 AA.
 AC P31953; P96806;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Antigen 85-C precursor (85C) (Antigen 85 complex C) (Ag85C) (Mycoly1
 DE transferase 85C) (EC 2.3.1.-) (Fibronectin-binding protein C).
 GN FBPC OR MP745 OR RV0129C OR MT0137 OR MTC15.03C.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Erdmann;
 RX MEDLINE=91348869; PubMed=1715324;
 RA Content J., la Cuvelier A., de Wit L., Vincent-Levy-Frebault V.,
 RA Ooms J., de Bruyn J.;
 RT "The genes coding for the antigen 85 complexes of Mycobacterium
 RT tuberculosis and Mycobacterium bovis BCG are members of a gene

RT family: cloning, sequence determination, and genomic organization of
RT the gene coding for antigen 85-C of M. tuberculosis.";
RN Infect. Immun. 59:3205-3212(1991).
RN [2]
RN SEQUENCE FROM N.A.

RP STRAIN=H37RV;
RC MEDLINE=98295987; PubMed=9634230;
RX Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekaita F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jajels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellon S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RA "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544(1998).
RL [3]

RN SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RT Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

RL [4]
RN X-RAY CRYSTALLOGRAPHY (1.50 ANGSTROMS).
RX MEDLINE=20122630; PubMed=10655617;
RA Renning D.R., Klabunde T., Besra G.S., Vissa V.D., Belisle J.T.,
RA Sacchettini J.C.;
RA "Crystal structure of the secreted form of antigen 85C reveals
RT potential targets for mycobacterial drugs and vaccines.";
RL Nat. Struct. Biol. 7:141-146(2000).
CC !- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
CC MYCOLITRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
CC FOR MAINTAINING CELL WALL INTEGRITY.

CC !- SUBCELLULAR LOCATION: Secreted.
CC !- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
CC N-TERMINAL OF C. GLUTAMICUM PSI PROTEIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC -----
CC EMBL; X57229; CAA40506.1; --
CC EMBL; Z92770; CAB07035.1; --
CC EMBL; AE006925; AAK44361.1; ALT_INIT.
CC PIR; D70615; D70615.
CC PDB; 1DQY; 12-JUL-00.
CC PDB; 1DQY; 12-JUL-00.
CC TIGR; MT0137; --
CC TubercuList; Rv0129c; --
CC InterPro; IPR008001; Esterase put.
CC Pfam; PF00756; Esterase; 1.
CC Transferase; Acyltransferase; Antigen; Signal; 3D-structure;
KW Complete proteome.
FT SIGNAL 1 46 POTENTIAL.
FT CHAIN 47 340 ANTIGEN 85-C.
FT ACT_SITE 170 170
FT ACT_SITE 274 274
FT ACT_SITE 306 306
FT ACT_SITE 21 21
FT CONFLICT 21 21 L -> V (IN REF. 1).
FT STRAND 54 61

65 62 65
66 66 66
73 73 73
79 79 79
82 82 82
86 86 86
87 87 87
94 94 94
98 98 98
101 101 101
105 105 105
107 107 107
111 111 111
115 115 115
120 120 120
123 123 123
125 125 125
134 134 134
138 138 138
140 140 140
141 141 141
146 146 146
147 147 147
157 157 157
169 169 169
170 170 170
171 171 171
182 182 182
184 184 184
186 186 186
189 189 189
193 193 193
199 199 199
200 200 200
202 202 202
203 203 203
214 214 214
215 215 215
216 216 216
220 220 220
224 224 224
228 228 228
230 230 230
234 234 234
235 235 235
240 240 240
241 241 241
247 247 247
248 248 248
250 250 250
254 254 254
263 263 263
268 268 268
291 291 291
292 292 292
296 296 296
308 308 308
317 317 317
318 318 318
319 319 319
327 327 327
340 340 340
AA; 36771 MW; 2D868C438D697988 CRC64;
SEQUENCE
Query Match 68.7%; Score 1198; DB 1; Length 340;
Best Local Similarity 71.6%; Pred. No. 2.5e-83;
Matches 222; Conservative 27; Mismatches 59; Indels 2; Gaps 1;
QY 13 RLIMICTAAAVLPGVLGLAGGAATAGAFRRPGLPVEYLQVPSMGRDIKVQFQSGGNN 72
Db 19 RLAIAMGAVLVYGLVGTGGPATAGAFRRPGLPVEYLQVPSMGRDIKVQFQGGPH 78
QY 73 SPAYLLDGLRAQDDYNGWDINTPAFEYYQSGLSIVMPVGGSSPYSDWSPACGKAC 132
Db 79 --AVILLDGLRAQDDYNGWDINTPAFEYYQSGLSIVMPVGGSSPYSDWSPACGKAC 136
QY 133 QYKWEFTLTSELQWLSANRAVKPTGSAAGLSMAGSSAMTILAAVHPOOFIYAGLSAL 192
Db 137 YTYKWEFTLTREMPAMLVQANKGVSPGTGNAAGVLSMAGSSAMTILAAVHPOOFIYAGLSAL 196
QY 193 LDPSQMGPSLLIGLAMDAGGYKAAADWGPSSDPAMERNDPTQCIIPKLVANNTRVWYCG 252
Db 197 LNPSEGWPVTLIGLAMDNDGGYNNANWGPSSDPAMERNDPTQCIIPKLVANNTRVWYCG 256
QY 253 NGTPNELGGANIPAEFLENFVRSNLTQDAYNAAGHNAVFPNPNTHSWYWGALIN 312
Db 257 NGTSPDLGGDNIPAKFLEGLTLTNTQTRDYAAGGRNGVFPNPNTHSWYWGALIN 316
QY 313 AMKGDLOSSL 322
Db 317 AMKADIQHLV 326
RESULT 14
A85C_MYCLE

ID AC Q05862; STANDARD; PRT; 333 AA.
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antigen 85-C precursor (85C) (Antigen 85 complex C) (Ag85C) (Mycoly1
 GN FBPC OR FBPC2 OR ML2655.
 OS Mycobacterium leprae.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CX NCBI_TaxID=1769;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93366419; PubMed=8359887;
 RA Rinke de Wit T.F., Bekelie S., Osland A., Wiele B.,
 RA Janson A.A.M., Thole J.E.R.;
 RT "The Mycobacterium leprae antigen 85 complex gene family:
 RT identification of the genes for the 85A, 85C, and related MPT51
 RT proteins.";
 RL Infect. Immun. 61:3642-3647(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX de Mendonca-Lima L.;
 RA Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TN;
 RX MEDLINE=21128732; PubMed=11234002;
 RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
 RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
 RA Murthy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
 RA Barrell B.G.;
 RT "Massive gene decay in the leprosy bacillus.";
 RL Nature 409:1007-1011(2001).
 CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
 CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
 CC MYCOLYTIC TRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
 CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
 CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
 CC N-TERMINAL OF C. GLUTAMICUM P51 PROTEIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; Z21951; CAA79949.1; -;
 DR EMBL; M90649; AAA91865.1; -;
 DR EMBL; AL583926; CAC32187.1; -;
 DR PIR; S32114; S32114.
 DR HSP; P31953; IDQY.
 DR Leproma; ML2655; -;
 DR InterPro; IPR000801; Esterase; 1.
 KW Pfam; PF00756; Esterase; 1.
 FT SIGNAL 1 46
 FT CHAIN 47 333
 FT ACT SITE 170 170
 FT ACT SITE 274 274 BY SIMILARITY.
 FT ACT SITE 306 306 BY SIMILARITY.
 SQ SEQUENCE 333 AA; 36392 MW; 50EADF1731E8EC2 CRC64;

Query Match 67.0%; Score 1170; DB 1; Length 333;
 Best Local Similarity 69.3%; Pred. No. 3.1e-81;
 Matches 214; Conservative 33; Mismatches 60; Indels 2; Gaps 1;
 QY 14 RLMGTAAAVVLPGLVGLAGGAATAGAFSRPGLPVEYLQVFPSPMGRDVKVQFQSGGNN 73
 Db |||||TALLAGLVGVGDTATAVAFSPKGLPVEYLQVFPSPMGRDVKVQFQSGGQH- 78
 QY 74 PAVYLLDGLRAQDDYNGWDINTPAFEVYQSGLSIVMPVGGQSSPYSWYSPACGACGQ 133
 Db |||||DGLRAQDDYNGWDINTPAFEVYHSGLSIVMPVGGQSSPYSWYSPQCGQGH 137
 QY 134 TYKWETFLTSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHPQOFLYAGSL 193
 Db |||||TFLTQEMPSLQANKVLTGNAVGLSSGSSALILASITYPOQFFYAAASL 197
 QY 194 DPSQMGPSLIGLAMGAGGYKAADWMPSSDPAWERNRDTQQIPKLVANNTRLVVYCGN 253
 Db |||||TFLTQEMPSLQANKVLTGNAVGLSSGSSALILASITYPOQFFYAAASL 197
 QY 198 NPSEGWPFTMIGLAMNDSSGYNANSMWGPSTDPKWKNDPMVQIPRLVANNTRLVVYCGN 257
 Db |||||TFLTQEMPSLQANKVLTGNAVGLSSGSSALILASITYPOQFFYAAASL 197
 QY 254 GTPNELGGANIPAEFFLEFVRSSNLKFDQDAYNAAGHNVAFFPPNGTHSWYWGAI 313
 Db |||||TFLTQEMPSLQANKVLTGNAVGLSSGSSALILASITYPOQFFYAAASL 197
 QY 314 MKGDLQSSL 322
 Db |||||TFLTQEMPSLQANKVLTGNAVGLSSGSSALILASITYPOQFFYAAASL 197
 QY 318 MKPDIIQIL 326
 Db |||||TFLTQEMPSLQANKVLTGNAVGLSSGSSALILASITYPOQFFYAAASL 197
 RESULT 15
 A85C MYCAV STANDARD; PRT; 352 AA.
 ID A85C MYCAV STANDARD; PRT; 352 AA.
 AC 052972;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Antigen 85-C precursor (85C) (Antigen 85 complex C) (Ag85C) (Mycoly1
 DE transferase 85C) (EC 2.3.1.-) (Fibronectin-binding protein C).
 GN FBPC.
 OS Mycobacterium avium.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 CX NCBI_TaxID=1764;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15769;
 RX MEDLINE=97427949; PubMed=9284137;
 RA Ohara N., Ohara-Wada N., Kitaura H., Nishiyama T., Matsumoto S.,
 RA Yamada T.;
 RT "Analysis of the genes encoding the antigen 85 complex and MPT51 from
 RT Mycobacterium avium.";
 RL Infect. Immun. 65:3680-3685(1997).
 CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
 CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
 CC MYCOLYTIC TRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
 CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
 CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
 CC N-TERMINAL OF C. GLUTAMICUM P51 PROTEIN.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; D87323; BAA24161.1; -;
 DR HSP; P31953; IDQY.
 DR InterPro; IPR000801; Esterase; 1.
 DR Pfam; PF00756; Esterase; 1.

entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

KW Transferase; Acyltransferase; Antigen; Signal.
FT SIGNAL 1 46 POTENTIAL.
FT CHAIN 47 352 ANTIGEN 85-C.
FT ACT SITE 170 170 BY SIMILARITY.
FT ACT SITE 274 274 BY SIMILARITY.
FT ACT SITE 306 306 BY SIMILARITY.
SQ SEQUENCE 352 AA; 37756 MW; A29277CB650D60A4 CRC64;

Query Match 66.8%; Score 1165; DB 1; Length 352;
Best local Similarity 66.4%; Pred. No. 7.9e-81;
Matches 215; Conservative 39; Mismatches 64; Indels 6; Gaps 2;

QY 6 RKIR----AWGRIMIGTAAVVLPLGLVLAGAAGTAGAFSPRLPYEYLOVPSMGRD 61
DB 8 RKLGRGAAATMPRLALAAVUGASLLSGVAVAGGSPVAGAFSKPLPYEYLEVPSPNGRN 67

QY 62 IKVQFGGNNFAVILLGLRADDYNGWDINTPAFEWYQSGSLSTVMPVGQSSPYSD 121
DB 68 IKVQFGGGPH--AVYLLDGLRAQDDYNGWDINTPAFEFYQSGLSVIMPVGQSSPYSN 125

QY 122 WSPACGKAGCQTYKWTFTLTSELPWLSANRAVKPTGSAAGISLMAGSAMILAAVHPQ 181
DB 126 WYFPSSNGQNYIKWTFITLTQEMPLWQSNKQVSPAGNAVGLSMGGGSAALILAAVYPQ 185

QY 182 QFIYAGSLALLDPSQGMPSLIGLAWGDAGGYKAADMWGPSSDPAPERNDPTQOIPLV 241
DB 186 QFYPAASLSGLFNPSEGWWPTLLIGLAWNDGSGYNANSMWGPSTDPAPKRNDDPMVQIPRLV 245

QY 242 ANNTRIWYCGNGTPELNGANIIPAEFLNFVRSNLLKFDAYNAAGHNAVFPPNGT 301
DB 246 ANNTRIWYCGNGTPELNGDNPVPAKFLGLETLRTNEQFNQNYAAAGRGVFPFANGT 305

QY 302 HSWYWGQAQLNAMKGDLOSSLGAG 325
DB 306 HSWYWNQQLNAMKPMQVLLSG 329

RESULT 16
A85A MYCWR
ID A85A MYCWR STANDARD; PRT; 139 AA.
AC Q9KH57;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Antigen 85-A (A85A) (Antigen 85 complex A) (A985A) (Mycolyl transferase
DE 85A) (EC 2.3.1.-) (Fibronectin-binding protein A) (Fragment).
GN FBPA.
OS Mycobacterium marinum.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1781;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NC275;
RA Stear T.P., Jenkin G.A., Johnson P.D.R., Davies J.K.;
RT "Comparative genetic analysis of Mycobacterium ulcerans and
RT Mycobacterium marinum reveals evidence of recent divergence.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: PROTEINS OF THE ANTIGEN 85 COMPLEX ARE RESPONSIBLE FOR
CC THE HIGH AFFINITY OF MYCOBACTERIA TO FIBRONECTIN. POSSESSES A
CC MYCOLYTIC TRANSFERASE ACTIVITY REQUIRED FOR THE BIOGENESIS OF
CC TREHALOSE DIMYCOLATE (CORD FACTOR), A DOMINANT STRUCTURE NECESSARY
CC FOR MAINTAINING CELL WALL INTEGRITY (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: CONTAINS ONE DISULFIDE BRIDGE (BY SIMILARITY).
CC -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE
CC N-TERMINAL OF C. GLUTAMICUM P51 PROTEIN.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

```

CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC
CC
CC
DR EMBL; AF271345; AAF86326.1; -.
DR HSSP; P31953; IDQZ.
DR InterPro; IPR00801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
KW Transferase; Acyltransferase; Antigen.
FT NON_TER 1 1 BY SIMILARITY.
FT ACT SITE 10 10 BY SIMILARITY.
FT ACT SITE 114 114 BY SIMILARITY.
FT NON_TER 139 139
SQ SEQUENCE 139 AA; 14449 MW; 6E0ED52C7CA48637 CRC64;

Query Match 35.8%; Score 625; DB 1; Length 139;
Best local Similarity 80.6%; Pred. No. 1.4e-40;
Matches 112; Conservative 14; Mismatches 13; Indels 0; Gaps 0;

QY 157 PTGSAAGISLMAGSAMILAAVHPQFIYAGSLALLDPSQGMPSLIGLAWGDAGGYKA 216
DB 1 PTGSGVGLSMAGSSALILAAVHPDQFVYSGSLALLDPSQGMPSLIGLAWGDAGGYKA 60

QY 217 ADMWGPSSDPAPERNDPTQOIPLVANNTLWYCGNGTENELGGANIPAEFLNFVRS 276
DB 61 SDMWGPKDDPAWRNDPMLQVGLVANNTRIWYCGNGKPSDLGGDNLPAKFLGFEVRS 120

QY 277 NLKFDQDAYNAAGHNAVFN 295
DB 121 NMRQAAVNAAGHNAVFN 139

RESULT 17
MP51 MYCTU
ID MP51 MYCTU STANDARD; PRT; 299 AA.
AC Q48923; Q33176;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE MPT51/MPB51 antigen precursor.
GN MPT51 OR MPB51 OR RV3803C OR MT3910 OR MTV026.08C.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773, 1765;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RA Oettinger T., Andersen P.;
RT "Evidence for the secreted protein MPT51 from Mycobacterium
RT tuberculosis is a T-cell antigen.";
RL Submitted (OCT-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Horsley T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=M.tuberculosis; STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,

```


Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A., Bishai W.;
 "Whole genome comparison of Mycobacterium tuberculosis clinical and laboratory strains";
 Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 [4]
 SEQUENCE FROM N.A.
 SPECIES=M.DOVIS; STRAIN=BCG / Tokyo;
 MEDLINE=95242057; PubMed=7725062;
 Ohara N., Kitauro H., Hotokezaka H., Nishiyama T., Wada N., Matsumoto S., Matsuo T., Naito M., Yamada T.;
 "Characterization of the gene encoding the MP951, one of the major secreted protein antigens of Mycobacterium bovis BCG, and identification of the secreted protein closely related to the fibronectin binding 85 complex";
 Scand. J. Immunol. 41:433-442(1995).
 -!- SUBCELLULAR LOCATION: Secreted.
 -!- SIMILARITY: TO OTHER MYCOBACTERIUM A85 ANTIGENS AND TO THE N-TERMINAL OF C. GLUTAMICUM PSI PROTEIN.

 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; AJ002150; CAA05211.1; -
 EMBL; AL022076; CAA17867.1; -
 EMBL; AE007185; AAK48276.1; ALT_INTT.
 EMBL; D26486; BAA05497.1; -
 PIR; G70887; G70887.
 HSP; P31953; IDQZ.
 TIGR; MT3910; -
 TubercuList; RV3803c; -
 InterPro; IPR00801; Esterase_put.
 Pfam; PF00756; Esterase; 1.
 Transferase; Acyltransferase; Signal; Antigen; Complete proteome.
 SIGNAL 1 33 POTENTIAL.
 FT CHAIN 34 299 MPT51/MPB51 ANTIGEN.
 FT CONFLICT 246 246 A -> T (IN REF. 1).
 FT SEQUENCE 299 AA; 31089 MW; 4E2E38F87AEDD73E CRC64;

```

Query Match      29.7%; Score 519; DB 1; Length 299;
Best Local Similarity 38.7%; Pred. No. 3.3e-32;
Matches 113; Conservative 39; Mismatches 124; Indels 16; Gaps 5;

QY 33 GGAATAGAFSRPGLPVEYLQVPSWGMGRDIKVFQSGGNNSPAYVLLDGLRAQDDYNGWD 92
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 21 GGVAAVAEPTAKAAYENLWYPSWGMGRDIPVAFLAGGPH--AYVLLDAFNAGPDVSNWV 78
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 93 INTPAFVWYQGLSIVMPVGGQSFFYSDWYSPACGKAGCOTYKWKETFLISELPQWL SAN 152
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 79 TAGNAMNTILAGKISVAVAPAGGAYSMYTNWEQDG-----SKQWDTFLSAELPDWTLAAN 131
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 153 RAVKVTGSAATGLSWAGSSMILAAHYHQOFIYAGSLISALIDPSQGMPSLIGLAMGDAG 212
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 132 RGLAFGGHAAVGAAGGGYGAMALAAAFHDPDRFGFAGSMGFLYPNTTNGAIAAGMQQFG 191
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 213 GYKAADMWGPSDPAWERNDTQOIPKLVANNTLWYCGNGTPNELGGANIPAEFL ENF 272
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 192 GVDYTNMGWAGPQLGRWKWHDPMVHASLLAQNTTRVWV-----SPTN--PGASDPAAMIGQA 246
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 273 VRS--SNLKFQDAYNAGCHNAVNFPPNGTHSWEYWGAIQINAMKGDLOSSL 322
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 AEAMGNRMFYNORYSGVGHNGHDFDPASGDNMGWSNAPQLGAMSGDIVCAI 298
    ||| : ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 18
DEPT5 MYCLE
IDT MPT5 MYCLE
AC Q05868; Q05207;
PRT; 301 AA.

```

01-NOV-1995 (Rel. 32, Created)
16-OCT-2001 (Rel. 40, Last sequence update)
16-OCT-2001 (Rel. 40, Last annotation update)
MPT51 antigen precursor.
MPT51 OR ML0098.
Mycobacterium leprae.
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
Corynebacteriineae; Mycobacteriaceae; Mycobacterium.
NCBI_TaxID=1769;
[1]
SEQUENCE FROM N.A.
STRAIN=Thai53;
Yin Y.;
"Studies of MPT51 like protein of Mycobacterium leprae.";
Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases.
[2]
SEQUENCE FROM N.A.
STRAIN=TN;
MEDLINE=21128732; PubMed=11234002;
Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,
Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
Barrell B.G.;
"Massive gene decay in the leprosy bacillus";
Nature 409:1007-1011(2001).
[3]
SEQUENCE OF 1-220 FROM N.A.
MEDLINE=93366419; PubMed=8359887;
Kinke de Wit T.F., Bekellie S., Osland A., Wiele B., Janson A.A.M.,
Thole J.E.R.;
"The Mycobacterium leprae antigen 85 complex gene family:
identification of the genes for the 85A, 85C, and related MPT51
proteins";
Infect. Immun. 61:3642-3647(1993).
-!- SUBCELLULAR LOCATION: Secreted.
-!- SIMILARITY: TO OTHER MYCOBACTERIUM 85 ANTIGENS AND TO THE
N-TERMINAL OF C.GLUTAMICUM PSI PROTEIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; D50488; BAA09079.1; -
EMBL; AL583917; CAC29606.1; -
EMBL; Z21949; CAA79947.1; -
DR PIR; B86921; B86921
DR PIR; S32111; S32111.
DR HSP; P31953; IDQZ.
DR Leproma; ML0098; -
DR InterPro; IPR00801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
KW Transferase; Acyltransferase; Signal; Antigen; Complete proteome.
FT SIGNAL; 1 36 POTENTIAL.
FT CHAIN 37 301 MPT51 ANTIGEN.
FT CONFLICT 41 41 S -> F (IN REF. 1).
FT CONFLICT 286 301 APOLAMSGDIVGAIR -> GAAVGCYVGRYRRHSLGRNQ
FT WHGRCCGDPVSCMGARWVSGEHRHRTPYVHRNLPRDLPI
FT LLAGDEPWGRKCHVSATSPSSGRLPCRWFPSRWPLMLW
FT (IN REF. 1).
SQ SEQUENCE 301 AA; 31191 MW; 3E6866C805BFAB28 CRC64;
Query Match 28.6%; Score 498.5; DB 1; Length 301;
Best Local Similarity 38.7%; Pred. No. 1.2e-30;
Matches 120; Conservative 39; Mismatches 124; Indels 27 Gaps 8;

QY 15 LMTGTAATVPLGLVAGGAATAGAFSPRLPVEYLQVPSMGRDVKIYQFOSGNNSP 74
 Db 16 LAVGVFAAALV--LAGTAGNKAAG-----YESLMVPSNAMGRDIPVAFMAGGPH-- 63
 QY 75 AVVILDLGRLAQDDYNGWDINTPAFEWYQSGLSITWMPVGGSSFYSDWSPACGKGCQT 134
 Db 64 AVVILDLAFNALDVSNWVTAGNMTLGGRGISVAPAGGAYSMTYTWNDG-----S 116
 QY 135 YKWTETITSELPQWLSANRAVKPTGSAAGISLWAGSSAMILAAYHPQOFTYAGSLALD 194
 Db 117 KQWDTFLUSSELPDWLAKRGAPDGHAAVAGSQGYAALAAAPHDRFGFAGSLSGFVY 176
 QY 195 PSQGMGFSLLGLAMGDAGGYKAAAMWGPSSDPWERNDDPTQOIPKLVANNTRLWVYCGNG 254
 Db 177 PSSTNYNGAILAGLQFGIDGNGMWGAPQLGRWKHDDPYVHSLAQNNTYRWVY---- 232
 QY 255 TPNELGANNIPAEFLNPFVRS--SNLKFQDAYNAGGHNAVFNPNGTHSWYWGAGQLN 312
 Db 233 SPMTMGG-DIDA-MIGQAVASMGSSREFYQYRSVGGHNGHDFDSGGDNGWGAWAPQLA 290
 QY 313 AMKGDLOSSL 322
 Db 291 AMSGDIVGAI 300

RESULT 19

ID CSP1_CORGL STANDARD; PRT; 657 AA.
 AC Q01377;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE PS1 protein precursor.
 GN CSP1 OR CGL2875.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 17965 / Melassecola;
 RX MEDLINE=93023863; PubMed=1406274;
 RA Joliff G., Mathieu L., Hahn V., Bayan N., Duchiron F., Renaud M.,
 RA Shechter E., Leblon G.;
 RT "Cloning and nucleotide sequence of the cspl gene encoding PS1, one
 RT of the two major secreted proteins of Corynebacterium glutamicum: the
 RT deduced N-terminal region of PS1 is similar to the Mycobacterium
 RT antigen 85 complex.";
 RL Mol. Microbiol. 6:2349-2362(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: ONE OF THE TWO MAJOR SECRETED PROTEINS.
 CC -!- SUBCELLULAR LOCATION: Secreted.
 CC -!- SIMILARITY: THE N-TERMINAL IS VERY SIMILAR TO THE COMPLETE
 CC SEQUENCES OF THE MYCOBACTERIAL ANTIGENS 85-A, 85-B & 85-C.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; X66078; CAA46877.1; .
 CC F01; S25184; S25184.
 CC HSSP; P31953; 1DQZ.

DR InterPro: IPR000801; Esterase_put.
 DR InterPro: IPR000379; Ser_estrs_site.
 DR Pfam: PF00756; Esterase; 1.
 KW Signal; Complete proteome; Pyrrolidone carboxylic acid.
 FT SIGNAL 1 43
 FT CHAIN 44 657 PS1 PROTEIN.
 FT MOJ RES 44 44 PYRROLIDONE CARBOXYLIC ACID (PROBABLE).
 FT CONFLICT 54 44 I -> V (IN REF. 1).
 FT CONFLICT 177 177 V -> I (IN REF. 1).
 FT CONFLICT 189 189 G -> E (IN REF. 1).
 FT CONFLICT 320 320 G -> D (IN REF. 1).
 FT CONFLICT 330 330 A -> T (IN REF. 1).
 FT CONFLICT 350 350 S -> N (IN REF. 1).
 FT CONFLICT 461 461 A -> D (IN REF. 1).
 FT CONFLICT 562 562 E -> D (IN REF. 1).
 FT CONFLICT 596 596 A -> V (IN REF. 1).
 SQ SEQUENCE 657 AA; 70674 MW; E2C3F1B57A7938A3 CRC64;
 Query Match 26.9%; Score 470; DB 1; Length 657;
 Best Local Similarity 36.1%; Pred. No. 4e-28;
 Matches 115; Conservative 40; Mismatches 140; Indels 24; Gaps 8;
 QY 20 AAAYVPLGLVGLAGGAATAGAFSPRLPVEYLQVPSM-GRDIKVV-----FQSGGN 71
 Db 78 AGEVTPETAGLPDGRVRVISAEWATSKIV-ILTQSAAMPERPIKVLQLLPRDWYSSNR 136
 QY 72 NSPAVYLDGLRAQDDYNGWDINTPAFEWYQSGLSITWMPVGGSSFYSDWSPACGKAG 131
 Db 137 EFPEIWDGLRLAIEEQSGWTIETIEQYADKNAIVLVPVGGSSFYSDWEGPNNGK-- 194
 QY 132 CQYKWTETITSELPQWLSANRAVKPTGSAAGISLWAGSSAMILAAYHPQOFTYAGSLA 191
 Db 195 --NYQWETFLTQELAPILDKGFR-SNTDRAITGISMGGTAAVNIATHHPDMKFKVGSFSG 251
 QY 192 LLDPSQGMGFSLLGLAMGDAGGYKAAAMWGPSSDPWERNDDPTQOIPKLVANNTRLWVYC 251
 Db 252 YLDTTSAGMPIAISALADAGGYDANAMWGPVGSERWQENDPKSNVDKL--KGKTIYSS 309
 QY 252 GNGTNP-----ELGGANIPAEFLNPFVRSNNLKFQDAYNAGGHNAVFNPNGTHSW 304
 Db 310 GNGADDFGKGSVAIGPANAAAGVLEVISMTSQTFVDRASQA-GVEVVASFRPFGVHSW 368
 QY 305 EYWGAGLQNLAMKGDLOSSLG 323
 Db 369 EYVQFEMTQAFPHIANALG 387
 PRT; 456 AA.
 RESULT 20
 YC88 MYCTU STANDARD; PRT; 456 AA.
 AC Q10614;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein Rvi1288.
 GN RVI1288 OR MT1326 OR WTCY373.07.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37Rv;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the

```

RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.F., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- DOMAIN: LYSM REPEATS ARE THOUGHT TO BE INVOLVED IN PPTIDOGLYCAN
CC BINDING.
CC -!- SIMILARITY: SOME, TO MYCOBACTERIUM A85 ANTIGENS.
CC -!- SIMILARITY: Contains 3 LysM repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; Z73419; CAA97754.1; -.
CC DR EMBL; AE007007; AAK45587.1; -.
CC DR PIR; D70772; D70772.
CC DR TIGR; MT1326;
CC DR Tuberculist; RVI288; -.
CC DR InterPro; IPR000801; Esterase_put.
CC DR InterPro; IPR002482; LysM.
CC DR InterPro; IPR003379; Ser estrs_site.
CC DR Pfam; PF00756; Esterase; 1.
CC DR Pfam; PF01476; LysM; 3.
CC DR SMART; SM00257; LysM; 3.
CC DR Hypothetical protein; Repeat; Complete proteome.
CC KW REPEAT 5 51 LYSM 1.
CC FT REPEAT 56 102 LYSM 2.
CC FT REPEAT 107 153 LYSM 3.
CC SQ SEQUENCE 456 AA; 49618 MW; 9870C0471EDF113A CRC64;
Query Match 8.4%; Score 146.5; DB 1; Length 456;
Best Local Similarity 25.7%; Pred. No. 0.00065;
Matches 68; Conservative 34; Mismatches 86; Indels 77; Gaps 16;
QY 66 FQSGNNSPAVYLLDGLRAQDDYNGWDI-----NTPAFEWYQSGLSIVMPVGQS 117
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
198 YRTSGRTYVLYLFGGGTQDQDFTDFLGRDLTAGKP-----IIIVPDGCHAG 248
QY 118 FYSDDWYSPACGKAGCQTYKWTFTLSELPOWLSAN--RAVKPTGSAAGLSMAGSSAMIL 175
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
249 WYSNPVSVFVGR-----NWETHIAQLPWIEANFRYAEYDGRAVGFSGMGFGALKY 303
QY 176 AAYHPQOFTYAGSLSDLPSCMGPSLI-----GL-----AMGDAGGYKAADWNGP 222
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
304 AAKYGHFASASHS-----GPASRRDFGLVHVHWNISAVLDLGG---GTVYGA 351
QY 223 SSDPAWER-----NDPTQIPKLVANTRLYWCNGN-----TPNE--LGGANIFA 266
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
352 ---PLMDQARVADNVERIDSY--RNKRIFLVAGTSPDPANWFDVNETQVLAGOR---- 403
QY 267 EFLENFVRSSNLKF-QDAYNAGGH 290
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
404 EPRE---LUSNAGIPHESVEFGH 425
RESULT 21
ID ESTD HUMAN STANDARD; PRT; 282 AA.
AC P10768;
DT 01-JUL-1989 (Rel. 11, Created)

```

```

DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Esterase D (EC 3.1.1.1).
GN ESD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88272298; PubMed=3164702;
RA Young L.-J.S., Lee E.Y.-H.P., To H., Bookstein R., Shew J.-Y.,
RA Donoso L.A., Sery T., Giblin M., Shields J.A., Lee W.-H.;
RT "Human esterase D gene: complete cDNA sequence, genomic structure,
RT and application in the genetic diagnosis of human retinoblastoma.";
RL Hum. Genet. 79:137-141(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Hypothalamus;
RX MEDLINE=20402571; PubMed=10931946;
RA Hu R.-M., Han Z.-G., Song H.-D., Peng Y.-D., Huang Q.-H., Ren S.-X.,
RA Gu Y.-J., Huang C.-H., Li Y.-B., Jiang C.-L., Fu G., Zhang Q.-H.,
RA Gu B.-W., Dai M., Mao Y.-F., Gao G.-F., Zhang C.-K., Ye M., Zhou J.,
RA Xu S.-H., Gu J., Shi J.-X., Jin W.-R., Zhong R.-K., Wu T.-M.,
RA Huang G.-Y., Chen Z., Chen M.-D., Chen J.-L.;
RT "Gene expression profiling in the human hypothalamus-pituitary-adrenal
RT axis and full-length cDNA cloning.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:9543-9548(2000).
RN [3]
RP PRELIMINARY SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=86313576; PubMed=3462698;
RA Lee E.Y.-H.P., Lee W.-H.;
RT "Molecular cloning of the human esterase D gene, a genetic marker of
RT retinoblastoma.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:6337-6341(1986).
RN [4]
RP SEQUENCE OF 150-186 FROM N.A., AND SEQUENCE OF 150-175.
RX MEDLINE=86313620; PubMed=3462714;
RA Squire J., Dryja T.P., Dunn J., Goddard A., Hofmann T.,
RA Musarella M., Willard H.F., Becker A.J., Gallie B.L., Phillips R.A.;
RT "Cloning of the esterase D gene: a polymorphic gene probe closely
RT linked to the retinoblastoma locus on chromosome 13.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:6573-6577(1986).
RN [5]
RP SEQUENCE OF 168-200 FROM N.A.
RA Tsuchida S., Ikemoto S., Kajii E.;
RX Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
RN [6]
RP VARIANT GLU-190.
RX MEDLINE=94171223; PubMed=7907313;
RA Tsuchida S., Fukui E., Ikemoto S.;
RT "Molecular analysis of esterase D polymorphism.";
RL Hum. Genet. 93:255-258(1994).
CC -!- CATALYTIC ACTIVITY: A carboxylic ester + H(2)O = an alcohol + a
CC carboxylic anion.
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC VESICLES.
CC -!- POLYMORPHISM: THERE ARE TWO MAJOR ELECTROPHORETIC ISOTYPES. THE
CC SEQUENCE OF THE ESD 1 VARIANT IS SHOWN.
CC -!- SIMILARITY: STRONG, TO YEAST HRE299.
CC -!- SIMILARITY: VERY LOW, TO TYPE-B CARBOXYLESTERASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M13450; AAA52408.1; ALT SEQ.
CC DR EMBL; AF112219; AAC99788.1; -.
CC DR EMBL; AF052509; AAC06298.1; -.

```

CC PIR; A23543; A23543.
 CC Genew; HGNC:3465; ESD.
 CC MIM; 133280;
 CC GO; 0016023; C:cytoplasmic vesicle; NAS.
 CC GO; 0004759; F:serine esterase activity; NAS.
 CC InterPro; IPR000801; Esterase_put.
 CC InterPro; IPR000379; Ser_estrs_site.
 CC Pfam; PF00756; Esterase; 1.
 CC Hydrolase; Serine esterase; Polymorphism.
 CC ACT SITE 149 149 POTENTIAL.
 CC ACT SITE 190 190 G -> E (IN ESD 2; dbSNP:1051064).
 CC VARIANT 190 190 /FTID=VAR_005202.
 CC SEQUENCE 282 AA; 31463 MW; BFC20D5FA2BBDCE CRC64;
 CC
 CC Query Match 7.2%; Score 125; DB 1; Length 282;
 CC Best Local Similarity 21.2%; Pred. No. 0.016;
 CC Matches 52; Conservative 38; Mismatches 91; Indels 64; Gaps 9;
 CC
 CC 74 PAVYLLDGLRAQDYNDINTPAFWYQS-----GLSIVMP-----V 112
 CC 46 PALLYWLSGLTCTEQ-----NFIKSGVGHQSAHEGLVVIAPDTPRCNKGEDSWDF 99
 CC DB
 CC 113 GQSSFSYSD-----WYSPACGKAGCQTYKWTETLTSELPQWLSANRAYKPTGSAIGLUM 167
 CC 100 GTGAGFYVDATEDPKWT-----NYPMYSYVTEELPOLINAFVDPQRMSTFGHSM 150
 CC 168 AGSSAMILAAVHPQOQFIYAGLSLALDPSQGMGSLIGLMDGAGGYKAADWGPSSDPA 227
 CC 151 GGHGALICALKNPGKYSVAFAPICNPV-----LCP-----WGKFAFGYLGTDSK 198
 CC 228 WERNPTQOQIPKLIVANNTRLVYCGNTPNELGGANIPAEFLENFVRSSN-----LKF 280
 CC 199 WKAYDATHLVKSPGSQDLIDLDQGDQDQFLDGLLP-----DNFIACTEKKIPVPRFL 254
 CC 281 QDAYN 285
 CC 255 QEGYD 259
 CC
 CC RESULT 22
 CC AGLE RHIME STANDARD; PRT; 458 AA.
 CC ID XANP XANS2
 CC AC Q60106;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE alpha-glucosidase-binding periplasmic protein agle precursor.
 CC AGLE OR R06995 OR SMC03061.
 CC GN Rhizobium meliloti (Sinorhizobium meliloti).
 CC OS Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
 CC OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
 CC NCBI_TaxID=382;
 CC RN
 CC SEQUENCE FROM N.A.
 CC RP MEDLINE=99328961; PubMed=10400573;
 CC RX Willis L.B., Walker G.C.;
 CC RA "A novel Sinorhizobium meliloti operon encodes an alpha-glucosidase
 CC and a periplasmic-binding-protein-dependent transport system for
 CC alpha-glucosides."
 CC RT J. Bacteriol. 181:4176-4184 (1999).
 CC RL [2]
 CC RN
 CC SEQUENCE FROM N.A.
 CC RP STRAIN=1021;
 CC RC MEDLINE=21396507; PubMed=11481430;
 CC RX Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,
 CC RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,
 CC RA Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
 CC RA Pohl T., Portetelle D., Puchler A., Purnelle B., Ramsperger U.,
 CC RA Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;
 CC RT "Analysis of the chromosome sequence of the legume symbiont
 CC Sinorhizobium meliloti strain 1021."
 CC RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882 (2001).
 CC RL
 CC -!- FUNCTION: PART OF THE BINDING-PROTEIN-DEPENDENT TRANSPORT SYSTEM

CC FOR ALPHA-GLUCOSIDES SUCH AS SUCROSE, MALTOSE AND TREHALOSE.
 CC -!- SUBCELLULAR LOCATION: Periplasmic (Probable).
 CC -!- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING
 CC PROTEIN FAMILY 1.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AF045609; AAD12050.1; ALT_INIT.
 CC EMBL; AL591784; CAC45267.1; -
 CC InterPro; IPR006061; SBP_dom1.
 CC Pfam; PF01547; SBP_bac_1; 1.
 CC PROSITE; PS01037; SBP_BACTERIAL_1; FALSE NEG.
 CC Sugar transport; Transport; Periplasmic; Signal; Complete proteome.
 CC SIGNAL 1 27 POTENTIAL.
 CC CHAIN 28 458 ALPHA-GLUCOSIDES-BINDING PERIPLASMIC
 CC PROTEIN AGLE.
 CC SEQUENCE 458 AA; 49703 MW; E7A7F8157C2FC291 CRC64;
 CC
 CC Query Match 6.6%; Score 116; DB 1; Length 458;
 CC Best Local Similarity 23.1%; Pred. No. 0.13;
 CC Matches 83; Conservative 27; Mismatches 118; Indels 132; Gaps 17;
 CC
 CC 13 RLRLMTGTAAVVLPGLVGLAGGATAGAFSRPG-----LPVEYLQVP 54
 CC 3 RSLLLGVAAFAALLAGTAGLAGTAGADLKFKFGEDSRFNWASLEEFKKGHLKGQTLTIF 62
 CC 55 SPSMGKD-----IKVQFSQGNNSPAVYLLDGLRAQDYNDINTPAF 98
 CC 63 GPWRGEDEALFKSVYAFVEATGVKELKYSSENVEQIWI-----DTQAG---SPDV 112
 CC 99 EMYQSGL-----SIVMPVQGSFYSYSDWSPACGKAGCQTYKWTETLTSELPQWLSA 151
 CC 113 AILPQPLIADLAARGLLTPLG-----DETQWLID 143
 CC 152 NEAVKPTGSAATGLS-----MAGSSAMI-----LAAYHPOQFIYAG-----SL 189
 CC 144 NYA---AGQSWDLSTYNGKDTGTSALYAFYKIDVKLSVWYVPENEDAGYEYKPTMEEL 200
 CC 190 SALLD--PSQGMGSLIGLMDGAGGYKAADWGPSSDPAWERNPTQOQIPKLIVANNTR 247
 CC 201 KALTEKIAEDGKPCWICIGLGGCATGWPAID-W---VEDIMLRTPQAFYDKWVKNEIPF 256
 CC 248 WYVCGMTPELGGANIPAEFLENFVRSSNLKFDQAYNAAGGHNAV-----FNPPPPGTHS 303
 CC 257 TDAAVTGALEFGW-----FARN-----DAF-VDGGAANAVASTDRSPKGLFS 299
 CC
 CC RESULT 23
 CC XANP XANS2 STANDARD; PRT; 827 AA.
 CC AC Q60106;
 CC DT 16-OCT-2001 (Rel. 40, Created)
 CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
 CC DT 28-FEB-2003 (Rel. 41, Last annotation update)
 CC DE Xanthomonas pepsin precursor (EC 3.4.21.101) (Xanthomonas aspartic
 CC proteinase) (Carboxyl proteinase) (XCP).
 CC OS Xanthomonas sp. (strain T-22).
 CC OC Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
 CC OC Xanthomonadaceae; Xanthomonas.
 CC NCBI_TaxID=136420;
 CC [1]
 CC SEQUENCE FROM N.A., AND CHARACTERIZATION.
 CC MEDLINE=97058302; PubMed=8902622;
 CC Oda K., Ito M., Uchida K., Shibano Y., Pukhara K.-I., Takahashi S.;
 CC "Cloning and expression of an isovaleryl peptstatin-insensitive
 CC carboxyl proteinase gene from Xanthomonas sp. T-22."
 CC J. Biochem. 120:564-572 (1996).
 CC RL

[2]
RP MUTAGENESIS OF ASPARTIC ACID AND GLUTAMIC ACID RESIDUES, AND ACTIVE
RP SITES.
RX MEDLINE=99419069; PubMed=10488127;
RA Oyama H., Abe S.-I., Ushiyama S., Takahashi S., Oda K.;
RT "Identification of catalytic residues of ppstatin-insensitive
RT carboxyl proteinases from prokaryotes by site-directed mutagenesis.";
RL J. Biol. Chem. 274:27815-27822(1999).
CC -!- CATALYTIC ACTIVITY: Cleavage of casein.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- PTM: AUTOCATALYTICALLY PROCESSED.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S53.
CC -!- SIMILARITY: Contains 1 PKD domain.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).

DR	EMBL: D83740; BAA12093.1; --	
DR	MESOPS, S53.002; --	
DR	InterPro: IPR000601; PKD_domain.	
DR	Pfam: PF00082; Peptidase_S8; 1.	
DR	Pfam: PF00801; PKD; 1.	
DR	Pfam: PF04151; PPC; 1.	
DR	Pfam: SMART; SM00089; PKD; 1.	
DR	PROSITE; PS50093; PKD; 1.	
KW	Hydrolase; Protease; Serine protease; Zymogen; Signal.	
FT	SIGNAL	1
FT	PROPEP	24
FT	CHAIN	238
FT	PROPEP	636
FT	DOMAIN	635
FT	ACT_SITE	406
FT	ACT_SITE	544
FT	ACT_SITE	585
FT	SEQUENCE	827 AA; 83706 MW; 21A333C4C683DBC8F CRC64;

```

Query Match      6.2%; Score 109; DB 1; Length 827;
Best Local Similarity 22.1%; Pred. No. 0.87;
Matches 63; Conservative 46; Mismatches 114; Indels 62; Gaps 13;

QY      28 LVGLAGCA-----ATGAFSRPGLVEYLQVPSPGRGDIKVFQSGNNSPA 75
Db      320 IVGIAGGVKOLIIFYTSANGDSSSGITDAGITASYNRAVTDNIAKLVN--SLGEDETA 376

QY      76 VYLLDGLRAODDYNQWDINTPAFEWYQSGLSIVMPFGGSSFYSDWYSPACKAGCQTY 135
Db      377 AQ-QSGTQAADDA-----IFQQA-----QGQT-----FSIASGDAG-VY 410

QY      136 KWTFYITSELPQWLSANRA-----VKPTGSAAILGLSAGSSAMLLAAHYHQQFI 184
Db      411 QWSTDTPTSGSPGVV-ANSAGTVKIDLTHYSVSPASPVIYQVGGTTLTSGTWTSGETV 469

QY      185 YAGSLCALIDPSQMGPSLIGLAMDAGGYKAADMGPSSDPAWERNDPTQQIKPLVANN 244
Db      470 WNEGLSAIA-PSQGDNNQRLWATGGGVSLYEAAAPSQSSVSSSTKRVGPDLPDAFDAASSG 528

QY      245 TRLVYVCGNGTNPNEGANIPAE-FLENFVRSSNLKFDQAYNAAQ 288
Db      529 ALIVV--NGSTQVCGTSLASPLFFVGAFARIESA---ANNAIG 566

```

RESULT 24		
ID	PHAL_PSELE	
AC	PHAL_PSELE	STANDARD;
AC	F52090;	PRT; 414 AA.
DT	01-OCT-1996	(Rel. 34, Created)
DT	01-OCT-1996	(Rel. 34, Last sequence update)
DT	28-FEB-2003	(Rel. 41, Last annotation update)

DE	Poly(3-hydroxyalkanoate) depolymerase C precursor (EC 3.1.1.-) (PHA
DE	depolymerase) (PHB depolymerase).
GN	PHAZ1.
OS	Pseudomonas lemoignei.
OC	Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC	Oxalobacteraceae; Paucimonas.
OX	NCBI_TaxID=29443;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=94094865; PubMed=8269961;
RA	Jendrossek D., Mueller B., Schlegel G.;
RT	"Cloning and characterization of the poly(hydroxyalkanoic acid)-
RT	depolymerase gene locus, phaz1, of Pseudomonas lemoignei and its gene
RT	product.";
RL	Eur. J. Biochem. 218:701-710(1993).
CC	-!- FUNCTION: SPECIFIC FOR POLY(HYDROXYALKANOIC ACID) CONSISTING OF
CC	MONOMERS OF FOUR OR FIVE CARBON ATOMS AND FOR P-
CC	NITROPHENYLBUTYRATE AS SUBSTRATES.
CC	-!- SUBCELLULAR LOCATION: secreted.
CC	-!- SIMILARITY: BELONGS TO THE ABHYDROLASE SUPERFAMILY. LIPASE FAMILY

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).

DR	EMEL; Z22595; CAA80310.1; --
DR	PIR; S39530; S39530.
DR	InterPro; IPR000379; Ser esters site.
KW	Hydrolase; Lipid degradation; Signal; Multigene family.
FT	SIGNAL 1 37
FT	CHAIN 38 414
FT	ACT SITE 154 154
FT	DOMAIN 329 332
FT	DOMAIN 333 338
FT	POLY-THR. 343 346
FT	DOMAIN 349 356
FT	POLY-THR.
SQ	SEQUENCE 414 AA; C143496 MW; C1F7CDQFC59FB111 CRC64;

Query Match	6.1%;	Score 106;	DB 1;	Length 414;
Best Local Similarity	24.2%;	Pred. No. 0.67;		
Matches	60;	Conservative 26;	Mismatches 96;	Indels 66;
				Gaps 13;

QY	72	NSPAVILDG----	LRQDDYNGDI--	NTPAFWYYQSL	IVMPVGQSGSF	DWYSPA	126
		: : :	: : :	: : :	: : :	: : :	
Db	64	NAPLVVALLHGCTQT	AAAEACWSALGN	THKFFVYV	POQS-----	GNSNKCWF	118
		: : :	: : :	: : :	: : :	: : :	
QY	127	CGKAGCQTKWTFIT	SELPOWLSANRAV	KPTGSAAIGLS	MAGSSAMTAA	AHHQQFTIYA	186
		: : :	: : :	: : :	: : :	: : :	
Db	119	DITRG-----	QGEALS	IKQVDMKMHAN	IDPSRVVVT	GLSAGFWTT	174
		: : :	: : :	: : :	: : :	: : :	
QY	187	GLSL-----	ALLDPSQGMPS	LIG-LAMGDAG	GYKA-----	ADMWGSSD--	225
		: : :	: : :	: : :	: : :	: : :	
Db	175	APIAGGPKCAT	MTSAPTCTMSP	GVDKTTPAAWG	DLARGSGYNG	PKKSIWHGSSDYT	234
		: : :	: : :	: : :	: : :	: : :	
QY	226	--PAWERNDPTQ	-----	IPKLIV--	ANNTRL-----	WVYCGNGT	256
		: : :	: : :	: : :	: : :	: : :	
Db	235	VAPA-NQNETVEQ	FTNYHIGIDQ	TPDVSDT	VGCFHKVYK	SANGTPLVET	293
		: : :	: : :	: : :	: : :	: : :	
QY	257	NELG-GAN	263				
		: : :	: : :				
Db	294	VDPGTGAN	301				
		: : :	: : :				

RESULT	25
GLII_HUMAN	
ID_GLII_HUMAN	STANDARD;
PRT;	1106 AA.
AC AC P8B1E1; Q8TDN9;	
DT DT 01-AUG-1988 (Rel. 08,	Created)
DD DD 01-AUG-1988 (Rel. 08,	Last sequence update)

Db 620 WRSRAEYPGVNPAGVTRRASDPAQAADRPAPARVQRFKSLGCVHTPTTVAGGGQNF--D 677

QY 122 WXPACGKAGCQTYKWTETLTSLPQLSANRAVKPTG---SAAIGLSWAGSAMLAAY 178

Db 678 PVLFP-----TSVYSPQPSITENAMDARGLQEEPEVGTSMVGSLNPFYDMF 724

QY 179 HQQFI-YAGSLSALLDPQSGMPSLIGLWAGDAGYKAADWMPSSDDAW-ERNPTQO 236

Db 725 PPTDTLGYGPGGAAPYAGRGPSLPLGPGPTNY-----GNPCQQAQSYDPDPTQE 778

QY 237 IPLKVVNTRLW-----VTCNGTNPGLGANIPAEFLNF 272

Db 779 -----TWGEFPFHSGLYPG---PKALGGTYSQCPLREHY 809

RESULT 26

TOB2 HUMAN

ID TOB2 HUMAN STANDARD; PRT; 344 AA.

AC Q14106; Q9BY97; Q9UBI0;

DT 01-NOV-1997 (Rel. 35, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE TOB2 protein (Transducer of erbB-2 2).

GN TOB2 OR TOB4 (OR KIAAL663).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

PX MEDLINE=20071127; PubMed=10602502;

RA Ikematsu N., Yoshida Y., Kawamura-Tsuzuku J., Ohsugi M., Onda M.,

RA Hirai M., Fujimoto J., Yamamoto T.;

RT "Tob2, a novel anti-proliferative Tob/BTG1 family member, associates

RT with a component of the CCR4 transcriptional regulatory complex

RT capable of binding cyclin-dependent kinases.";

PL Oncogene 18:7432-7441(1999).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=21156230; PubMed=11258795;

RA Hirosewa M., Nagase T., Muranashi Y., Kikuno R., Ohara O.;

RT "Identification of novel transcribed sequences on human chromosome 22

RT by expressed sequence tag mapping.";

RL DNA Res. 8:1-9(2001).

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE=20057165; PubMed=10591208;

RA Dunham I., Hunt A.R., Collins J.E., Bruskewich R., Beare D.M.,

RA Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,

RA Bagguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,

RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,

RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,

RA Clegg S.M., Copley V.B., Cole C.G., Collier R.E., Connor R.,

RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,

RA Evans K.L., Fey J.M., Fleming S.J., Durbin R.M., Ellington A.G.,

RA Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,

RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,

RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,

RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,

RA Martyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,

RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,

RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,

RA Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,

RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,

RA Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,

RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,

RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,

RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,

RA Minoshima S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,

RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,

RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,

RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.T.,

RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,

RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,

RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,

RA Zhang M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,

RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,

RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,

RA Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,

RA Scheet P., Walker C., Wamsley L., Wohldmann P., Pepin K., Nelson J.,

RA Korf I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,

RA Emanuel B.S., Shaikh T., Kurahashi H., Saitta S., Budarf M.L.,

RA McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,

RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyraud M., Kedra D.,

RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,

RA Wilkison P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,

RA Tilahun Y., Wright H.;

RT "The DNA sequence of human chromosome 22.";

RN Nature 402:489-495(1999).

RL [4]

RN SEQUENCE FROM N.A.

RP TISSUE=Uterus;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,

RA Raha S.S., Loquellano N.A., Peters G.J., Carninci P., Prange C.,

RA Bosak S.A., McRwan P.J., McKernan K.J., Abramson R.D., Mullahy S.J.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RT human and mouse cDNA sequences.";

RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -!- FUNCTION: Anti-proliferative protein inhibits cell cycle

CC progression from the G0/G1 to S phases.

CC -!- SUBUNIT: Associates with CAP1.

CC -!- SUBCELLULAR LOCATION: Cytoplasmic.

CC -!- TISSUE SPECIFICITY: Ubiquitous.

CC -!- SIMILARITY: BELONGS TO THE BTG FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC

DR EMBL; D64109; BAA10971.1; -

DR EMBL; AB035207; BAA87042.1; -

DR EMBL; AB051450; BAB33333.1; ALT_INIT.

DR EMBL; AL008582; CAB62938.1; -

DR EMBL; BC038957; AAH38957.1; -

DR MIM; 607396; -

DR GO; GO:0005634; Cnucleus; TAS.

DR GO; GO:0008285; P:negative regulation of cell proliferation; TAS.

DR GO; GO:0007292; P:ogenesis; TAS.

DR GO; GO:0000074; P:regulation of cell cycle; TAS.

DR InterPro; IPR002087; Anti-proliferat.

DR Pfam; PF01211; Anti-proliferat; 1.

DR PRINTS; PR00310; ANTIPLRFBTGL.

DR SMART; SM00099; btg1; 1.

DR PROSITE; PS00960; BTG 1; 1.

DR PROSITE; PS01203; BTG_2; 1.

FT CONFLICT 34 35 RL -> QA (IN REF. 1; BAA10971).
 FT CONFLICT 214 214 P -> H (IN REF. 1; BAA10971).
 FT CONFLICT 278 317 QSGTGPFGSGAGTGNSSFDMAQVFGGANSLELEKT
 FT -> RAAAPQARLEAVGLAPATAAALTPWRYLEVVPASSWR
 FT RH (IN REF. 1; BAA10971).
 FT SEQUENCE 344 AA; 36632 MW; ACE4CD8939641BD5 CRC64;
 Query Match 5.8%; Score 101; DB 1; Length 344;
 Best Local Similarity 24.4%; Pred. No. 1.3;
 Matches 67; Conservative 41; Mismatches 115; Indels 52; Gaps 12;
 QY 40 AFSPGLPVEYLQVPSR---SMGFD-LKQVQSGGNSPAVYLLDGLRLAODYNGWDINT 95
 Db 71 AAKSGGLAVEDVRANVPEESVWIDPEVSTQIGKGAVKVLYL-----DDSG--CGA 122
 QY 96 PAFWYQSGLS-----IWPVGGQSSFYSDWYSPACGKAGCQTY-----KWETFLTSLP 146
 Db 123 PELDKETSSNPDAQVFPVIGSQDSSLSNPSFSGSPSTFIPRSAQPIPTTASFA 182
 QY 147 QWLSANRAVKPTGSAALGLSAGSAMLAAHYHQQFYAGSLALLDPSQGMGSLIGL 206
 Db 183 ATKEGSTKMGKGAAGGAVSSGAG--GQQPPQPPRMARSPTNSLLKHKSLSLMSHSL 240
 QY 207 AMGAGGYKAADMWGPSSDPAWERNPTQQIPKLVAANTRLWYCGNCTPN-----ELG 260
 Db 241 NF-----ITANPA-----POSQ-----SPNAKEFVYNGGSPSLFFDAADGQ 278
 QY 261 GANIPAELENFVRSSN-LKFQDAYNAAGHNAVF 294
 Db 279 GSGTGPFGGSGAGTCNSSFDMAQVFGGANSLE 313
 RESULT 27
 Y967 CORGL STANDARD; PRT; 426 AA.
 ID Y967 CORGL
 AC P35866;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Hypothetical protein Cg10967 (ORF2).
 GN CGL0967.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RX MEDLINE=92236417; PubMed=1667221;
 RA Seep-Feidhaus A.H., Kalinowski J., Puehler A.;
 RT "Molecular analysis of the Corynebacterium glutamicum lys1 gene
 involved in lysine uptake."
 RL Mol. Microbiol. 5:2995-3005(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X60312; CAA42856.1; -
 CC EMBL; AP005277; BAB98360.1; -
 CC PIR; S28812; S28812.
 CC InterPro; IPR000801; Esterase_put.
 CC InterPro; IPR000379; Ser_estrs_site.

DR Pfam; PF00756; Esterase; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 426 AA; 45730 MW; F0482C5F681200D8 CRC64;
 Query Match 5.8%; Score 100.5; DB 1; Length 426;
 Best Local Similarity 21.5%; Pred. No. 1.8;
 Matches 69; Conservative 43; Mismatches 122; Indels 87; Gaps 16;
 QY 29 VLAGGA--ATAGAFRPG---LPVEYLQVPSRMDIKVQFQSGGNSPAVYLLDGLR 83
 Db 162 VPLAGTDDTSFGDARDAYAYIPAYWNPSSLQ-----PVLVMPGNP 206
 QY 84 AODYNGW-----DINTPAFETWYQGLSIVMPVGGQSSFYSDWYSPACGKAGCQTYK 137
 Db 207 GQPDQ--WFSSGNADQTNFQATHDGIPIVSDVTGTSFG---NPACVDSQAQVM- 260
 QY 138 EFTLTSLPQWLSAN-RAVPTGSAAL-GLSMAGSSAMILAAHYHQQF-----IYAGSLSA 191
 Db 261 -TYLSHDVPMILIKQFRVNDQRTWTIGLSYGTGTCALQIWNHPEAYGSLDFSGQEEP 319
 QY 192 LLDPSCQMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNPTQQIPKLVAANTRLWYVC 251
 Db 320 TLGTRQQTVDQLFQ---GDEDAFKAV-----NPEDLLNQAISS----- 354
 QY 252 GNGTPELGGANIPAELENFVRSSN-----LKQDAYNAAGHNAVFPPNGTSHW 304
 Db 355 -----GAHTYSGISGRFIAGSNKSAVSALSHLDNLSNQAGMSTTFDTVAGG-HSF 404
 QY 305 EYGAQQLN-----AMKGLQ 319
 Db 405 QWRVALANTFDWVAKRGGLQ 425
 RESULT 28
 CBPB CANFA STANDARD; PRT; 416 AA.
 ID CBPB CANFA
 AC P55261;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Carboxypeptidase B precursor (EC 3.4.17.2) (47 kDa zymogen granule
 membrane associated protein) (ZAP47).
 GN CPB1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RA Fukuoka S.-I.;
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
 CC -|- CATALYTIC ACTIVITY: Peptidyl-L-lysine(or L-arginine) + H(2)O =
 CC peptide + L-lysine(or L-arginine).
 CC -|- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M14.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D78348; BAA11366.1; -
 CC EMBL; P09955; IPHA.
 CC InterPro; IPR003146; Propap_M14.
 CC InterPro; IPR000834; Zn_carbopept.
 CC Pfam; PF02244; Propap_M14; 1.
 CC Pfam; PF00246; Zn_carbopept; 1.
 CC PRINTS; SM00765; CRBOXYPTASEA.
 CC SMART; SM00631; Zn_pept; 1.

Fri Feb 6 13:54:48 2004

```

RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP IDENTIFICATION.
RA Rudd K.E.;
RL Unpublished observations (MAR-1996).
CC -!- SIMILARITY: STRONG, TO E.COLI YEIG AND H.INFLUENZAE HI0184.
CC -!- SIMILARITY: STRONG, TO HUMAN ESTERASE D.
CC -!- CAUTION: REF.1 SEQUENCE DIFFERS FROM THAT SHOWN DUE TO A
CC FRAMESHIFT IN POSITION 137.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D85613; -; NOT ANNOTATED_CDS.
CC EMBL; AE000142; AAC73458.1; -.
CC EMBL; U73857; AAL18080.1; -.
CC PIR; C64763; C64763.
CC EcoGene; EG13295; Yalm.
CC InterPro; IPR00801; Esterase_put.
CC InterPro; IPR00379; Ser_estr_site.
CC Pfam; PF00756; Esterase; 1.
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 277 AA; 31424 MW; C69C708293FA0AEC CRC64;
SQ
Query Match 5.6%; Score 97.5; DB 1; Length 277;
Best Local Similarity 18.8%; Pred. No. 1.9; Indels 81; Gaps 11;
Matches 52; Conservative 42; Mismatches 101;
74 PAVYLDGLRAQDDYNGWDINTPAPEWYQSGLS-----IVMP----- 111
44 PVLYWLSGLTNEQ-----NFITKSGQRYAAEHNIIVVAPDTSRGRSHVADAD 92
112 ---VGQSSGF-----SDWYSPACGKAGCQTYKWTFTLTSELPQWL-----SANRAVKP 157
93 RYDLGGQAGFYLNATQAPWN-----EHYKMYDYIRNELPDLVWHHFPATAKSIS- 142
158 TGSAAITGLNAGSSAMILAAHYHQQFIYAGSLSALLDPSQCMGSPSLIGLAWGDAGGYKAA 217
143 -----GHSNGGLGALVLRNPDEYVSVSAFSPVSPSQ-----VPWGOQAF 184
218 DMWGPSSDPARERNPTQIQIPKLVANNTRL-WVYCGNGTPELGGANIPAEFTLENFVRSS 276
185 AAYLAENKDWLDYDPV-----SLISQQRVARINVDGLSDDFYAEQLRTPNLEKICQEM 240
277 NLKXFQDAYNAAGHNAVNFPPNGTHSWYWGQQLN 312
241 NIKTLIRYHGYDHSYFVSVSFIGEHI-AYHANKLN 275

```

Search completed: February 5, 2004, 17:38:51
Job time : 18.4762 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:34:22 ; Search time 47.2024 Seconds
(without alignments)
1776.756 Million cell updates/sec

Title: US-09-805-427A-2
Perfect score: 1745
Sequence: 1 MTDVSRKIRAWRRRLMIGTA.....YWGALNAMKGLDQLSSLGAG 325

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830521

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database :

SPTREMBL 23:*

- 1: sp archaea:*
- 2: sp bacteria:*
- 3: sp fungi:*
- 4: sp human:*
- 5: sp invertebrate:*
- 6: sp mammal:*
- 7: sp mhc:*
- 8: sp organelle:*
- 9: sp phage:*
- 10: sp plant:*
- 11: sp rodent:*
- 12: sp virus:*
- 13: sp vertebrate:*
- 14: sp unclassified:*
- 15: sp rvirus:*
- 16: sp bacteriaph:*
- 17: sp archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match	Length	DB	ID	Description
1	1528	87.6	330	2	Q8VU82		Q8VU82 mycobacteri
2	1515	86.8	330	2	Q8L371		Q8L371 mycobacteri
3	1162	66.6	352	2	Q8KY58		Q8KY58 mycobacteri
4	1106.5	63.4	349	2	Q8KY59		Q8KY59 mycobacteri
5	634	36.3	136	2	P71483		P71483 mycobacteri
6	621	35.6	139	2	Q9F7W5		Q9F7W5 mycobacteri
7	553	31.7	140	2	P71531		P71531 mycobacteri
8	544	31.2	136	2	P71519		P71519 mycobacteri
9	473	27.1	669	16	Q8FM00		Q8FM00 corynebacte
10	428.5	24.6	120	2	P72063		P72063 mycobacteri
11	415	23.8	341	16	Q8NL31		Q8NL31 corynebacte
12	404	23.2	360	16	Q8FLZ9		Q8FLZ9 corynebacte
13	346	19.8	96	2	P71536		P71536 mycobacteri
14	303.5	17.4	411	16	Q8NRL7		Q8NRL7 corynebacte
15	297.5	17.0	77	2	Q9R667		Q9R667 mycobacteri
16	284.5	16.3	390	16	Q8FT53		Q8FT53 corynebacte

17	279	16.0	462	16	Q8NRW9	Q8NRW9 corynebacte
18	277	15.9	483	16	Q8NNL2	Q8NNL2 corynebacte
19	263	15.1	358	2	Q93R90	Q93R90 corynebacte
20	260	14.9	52	2	Q9S6X0	Q9S6X0 mycobacteri
21	257.5	14.8	381	16	Q8FSM7	Q8FSM7 corynebacte
22	252	14.4	484	16	Q8FQY3	Q8FQY3 corynebacte
23	238.5	13.7	365	16	Q8NTG4	Q8NTG4 corynebacte
24	233.5	13.4	342	16	Q86543	Q86543 streptomyc
25	192	11.0	39	2	Q9R5J7	Q9R5J7 mycobacteri
26	167	9.6	50	2	Q93JZ6	Q93JZ6 mycobacteri
27	167	9.6	261	16	Q99ZY1	Q99ZY1 streptococc
28	166	9.5	261	16	Q8P189	Q8P189 streptococc
29	156	8.9	37	2	Q9R5J5	Q9R5J5 mycobacteri
30	144.5	8.3	94	16	Q8VKB0	Q8VKB0 mycobacteri
31	143	8.2	259	16	Q97S09	Q97S09 streptococc
32	143	8.2	259	16	Q8DQR2	Q8DQR2 streptococc
33	138	7.9	58	2	O52955	O52955 mycobacteri
34	136	7.8	276	16	Q8P5F4	Q8P5F4 xanthomonas
35	134	7.7	252	16	Q928K5	Q928K5 listeria in
36	133	7.6	252	16	Q8Y4K5	Q8Y4K5 listeria mo
37	132	7.6	258	16	Q9L9X0	Q9L9X0 lactococcus
38	129.5	7.4	76	16	P71563	P71563 mycobacteri
39	127	7.3	258	2	Q9RAU7	Q9RAU7 lactococcus
40	126	7.2	288	5	Q811B5	Q811B5 drosophila
41	125	7.2	377	16	Q9X8H6	Q9X8H6 streptomyc
42	123.5	7.1	282	4	Q9BVJ2	Q9BVJ2 homo sapien
43	123	7.0	263	16	Q8DTB4	Q8DTB4 streptococc
44	121.5	7.0	258	2	Q8KWI9	Q8KWI9 lactococcus
45	121.5	7.0	278	2	O52299	O52299 anabaena az
46	121	6.9	276	16	Q8PPF0	Q8PPF0 xanthomonas
47	119.5	6.8	269	2	Q9E2Z5	Q9E2Z5 pasteurella
48	118.5	6.8	258	2	Q9L9W2	Q9L9W2 lactococcus
49	118.5	6.8	258	2	Q8KXE1	Q8KXE1 lactococcus
50	117.5	6.7	269	16	Q9CKZ9	Q9CKZ9 pasteurella
51	114.5	6.6	806	3	Q9P4T8	Q9P4T8 agaricus bi
52	113	6.5	282	6	Q9GJT2	Q9GJT2 sus scrofa
53	113	6.5	282	11	Q9ROP3	Q9ROP3 mus musculu
54	113	6.5	282	16	Q8YTB5	Q8YTB5 anabaena sp
55	113	6.5	312	11	Q9CWI4	Q9CWI4 mus musculu
56	112	6.4	272	10	Q8LAS8	Q8LAS8 arabidopsis
57	112	6.4	284	10	O22215	O22215 arabidopsis
58	111	6.4	322	16	Q8UDW6	Q8UDW6 agrobacteri
59	111	6.4	618	16	Q9A3F8	Q9A3F8 caulobacter
60	110.5	6.3	262	16	Q8DZC4	Q8DZC4 streptococc
61	110	6.3	282	6	Q9GM98	Q9GM98 sus scrofa
62	109.5	6.3	262	16	Q8E4Y4	Q8E4Y4 streptococc
63	109.5	6.3	577	2	Q9RE16	Q9RE16 arthrobacte
64	109	6.2	363	16	Q9RK16	Q9RK16 streptomyc
65	108	6.2	252	16	Q8DF76	Q8DF76 vibrio vuln
66	108	6.2	264	2	O31030	O31030 pectobacter
67	108	6.2	276	16	P73137	P73137 synochocyst
68	107.5	6.2	286	5	Q9VDP9	Q9VDP9 drosophila
69	107.5	6.2	689	16	Q8YU28	Q8YU28 anabaena sp
70	106	6.1	20	2	Q9R5R7	Q9R5R7 mycobacteri
71	105.5	6.0	293	16	Q99RU5	Q99RU5 staphylococ
72	105.5	6.0	293	16	Q8NV67	Q8NV67 staphylococ
73	105.5	6.0	415	16	Q9WX03	Q9WX03 streptomyc
74	105	6.0	430	16	Q9I739	Q9I739 pseudomonas
75	104	6.0	578	2	Q8ZG19	Q8ZG19 yersinia pe
76	104	6.0	280	16	Q8GRC2	Q8GRC2 marinobacte
77	104	6.0	963	3	O43024	O43024 schizosacch
78	103.5	5.9	277	16	Q8X5U5	Q8X5U5 escherichia
79	103.5	5.9	279	16	Q8EPC6	Q8EPC6 shewanella
80	103.5	5.9	279	16	Q8DL13	Q8DL13 synochococc
81	103	5.9	274	2	Q8GR67	Q8GR67 abiotrophia
82	102.5	5.9	849	16	Q8PPH0	Q8PPH0 xanthomonas
83	101.5	5.8	277	16	Q92OE0	Q92OE0 rhizobium m
84	101.5	5.8	437	16	Q8DHQ9	Q8DHQ9 synochococc
85	101.5	5.8	510	5	O44018	O44018 leishmania
86	101	5.8	279	2	Q51671	Q51671 paracoccus
87	101	5.8	289	16	Q8Y1T5	Q8Y1T5 talstonia s
88	100.5	5.8	328	16	Q9A7U1	Q9A7U1 caulobacter
89	100	5.7	466	16	Q92LZ9	Q92LZ9 rhizobium m

082950 pseudomonas

90 99.5 5.7 576 2 082950

ALIGNMENTS

RESULT 1

Q8VU82 PRELIMINARY; PRT; 330 AA.
AC Q8VU82;
DT 01-MAR-2002 (TremBrel. 20, Created)
DT 01-MAR-2002 (TremBrel. 20, Last sequence update)
DT 01-JUN-2002 (TremBrel. 21, Last annotation update)
DE Fibronectin-binding antigen 85 complex B.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
SEQUENCE FROM N.A.
RA Chang Y.-F., Shin K.-S.;
RT "Cloning and sequencing of gene encoding the antigen 85 complex B from
RL Mycobacterium paratuberculosis."
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF219121; AAL56853.1; -. put.
DR InterPro; IPR000801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
SQ SEQUENCE 330 AA; 34693 MW; 7E5C474F40C01001 CRC64;

Query Match 87.6%; Score 1528; DB 2; Length 330;
Best Local Similarity 85.5%; Pred. No. 2.8e-108;
Matches 277; Conservative 22; Mismatches 25; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDLSEKVRAGRRLLVGAATAVTLPGLLGLAGGAATANAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSVMPVGGSSFY 120
DB 61 DIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSVMPVGGSSFY 120
QY 121 DWYSPACGKAGCTYKWTETLTSELPWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
DB 121 DWYSPACGKAGCTYKWTETLTSELPWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
QY 181 QQFIYAGLSALLDPSQGMGSLIGLAWDAGGYKAADMWGPSSDPWERNPTQOIPKL 240
DB 181 DQFIYAGLSALLDPSQGMGSLIGLAWDAGGYKADAMWGPSSDPWQRNDPSLHPEL 240
QY 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFQDAYNAAGGHNAVENFPNG 300
DB 241 VGHNTRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFQDAYNGAGGHNAVENFNANG 300
QY 301 THSWEYGAQLNAMKPDLOSSLGA 324
DB 301 THSWEYGAQLNAMKPDLOSSLGA 324

RESULT 2

Q8L371 PRELIMINARY; PRT; 330 AA.
AC Q8L371;
DT 01-OCT-2002 (TremBrel. 22, Created)
DT 01-OCT-2002 (TremBrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBrel. 23, Last annotation update)
DE 85B protein.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19698;

RX MEDLINE=22000114; PubMed=12005331;
RA Mullerad J., Bercovier H.;
RT "The immunogenicity of Mycobacterium paratuberculosis 85B antigen."
RL Med. Microbiol. Immunol. 190:179-184 (2002).
DR EMBL; AF389346; AAM21939.1; -.
SQ SEQUENCE 330 AA; 34823 MW; 00F8D0CF079084A8 CRC64;

Query Match 86.8%; Score 1515; DB 2; Length 330;
Best Local Similarity 84.9%; Pred. No. 2.7e-107;
Matches 275; Conservative 22; Mismatches 27; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDLSEKVRAGRRLLVGAATAVTLPGLLGLAGGAATANAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSVMPVGGSSFY 120
DB 61 DIKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSVMPVGGSSFY 120
QY 121 DWYSPACGKAGCTYKWTETLTSELPWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
DB 121 DWYSPACGKAGCTYKWTETLTSELPWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
QY 181 QQFIYAGLSALLDPSQGMGSLIGLAWDAGGYKAADMWGPSSDPWERNPTQOIPKL 240
DB 181 DQFIYAGLSALLDPSQGMGSLIGLAWDAGGYKADAMWGPSSDPWQRNDPSLHPEL 240
QY 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFQDAYNAAGGHNAVENFPNG 300
DB 241 VGHNTRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFQDAYNGAGGHNAVENFNANG 300
QY 301 THSWEYGAQLNAMKPDLOSSLGA 324
DB 301 THSWEYGAQLNAMKPDLOSSLGA 324

RESULT 3

Q8KY58 PRELIMINARY; PRT; 352 AA.
AC Q8KY58;
DT 01-OCT-2002 (TremBrel. 22, Created)
DT 01-OCT-2002 (TremBrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBrel. 23, Last annotation update)
DE Antigen 85C complex.
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RA Chang Y.-F., Shin K.-S.;
RT "Cloning and characterization of Mycobacterium avium subsp.
RL paratuberculosis gene 85C."
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF280068; AAM73518.1; -.
SQ SEQUENCE 352 AA; 37754 MW; 39FFD0D7183975A8 CRC64;

Query Match 66.6%; Score 1162; DB 2; Length 352;
Best Local Similarity 66.0%; Pred. No. 2.1e-80;
Matches 214; Conservative 40; Mismatches 64; Indels 6; Gaps 2;

QY 6 RKIR----AWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGRD 61
DB 8 RKLRAAATMPERLAIAVAGSLLSGVAAGGSPVAGAFSKPGLPVEYLEVPSMGRN 67
QY 62 IKVFQSGGNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSVMPVGGSSFYSD 121
DB 68 IKVFQGGGPH--AVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSVMPVGGSSFY 125
QY 122 WYSPACGKAGCTYKWTETLTSELPWLSANRAVKPTGSAALGLSMAGSSAMILAAHPQ 181
DB 126 WYQSSGNGQNTYKWTETLTQEMPLWQSNQVSPAGNAVGLSMGSSALIAAYYPQ 185

QY 182 QTYIAGSLSDLPSCMGPSLIGLAWDAGGYKAADWGPSSDPAWERNNDPTQIPLV 241
 Db 186 QFYAASLGSFNPSEGWPTLLGLANDSGGNANSMWGPSTDPKAKRNDPMVQIPRV 245
 QY 242 ANNTLWVYCGNGTNPGLGAGNIPAFLENFVRSSNLKFDQAYNAAGGHNAVFPNPGT 301
 Db 246 ANNTLWVYCGNGTNPGLGAGNIPAFLENFVRSSNLKFDQAYNAAGGHNAVFPNPGT 305
 QY 302 HSWYVWGAQLNNAKMGDLOSLGAG 325
 Db 306 HSWPYWQQMAKMPDMQVLLUG 329

RESULT 4
 Q8KY59 PRELIMINARY; PRT; 349 AA.
 ID Q8KY59
 AC Q8KY59
 DT 01-OCT-2002 (TremBLrel. 22, Created)
 DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
 DE 01-MAR-2003 (TremBLrel. 23, Last annotation update)
 OS Antigen 85A complex.
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1770;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chang Y.-F., Shin K.-S.;
 RT "Cloning and characterization of Mycobacterium avium subsp.
 paratuberculosis gene for 85A."
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF280067; AAM73517.1; --
 SQ SEQUENCE 349 AA; 36763 MW; E3EB5DDDDACCSC085 CRC64;

Query Match 63.4%; Score 1106.5; DB 2; Length 349;
 Best Local Similarity 64.0%; Pred. No. 3.4e-76;
 Matches 219; Conservative 34; Mismatches 58; Indels 31; Gaps 7;

QY 1 MTDVSR---KIRAWGRRLMTGTAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVSPS 57
 Db 1 MTLVDRLGAVAGMPRLVVGAGALLSLGLGAVGSGATAGAFSRPGLPVEYLQVPSAA 60
 QY 58 MGRDIKVFQSGGNNPVALYLLDLRLAQDDYNGWDINTPAFWYQSGLSIVMPVGGSS 117
 Db 61 MGRDIKVFQSGGANGPALYLLDGMRAQDDFNWDINTPAFWYQSGLSIVMPVGGQSS 120
 QY 118 FVSDWYSPACGKAGCQTYKWTFTLSELPWLSANRAVKPTGSAALGLSMAGSSAMIL-- 175
 Db 121 FVSDWYKPCGKAGCTYKWTFTLSELPWLSAQKQVKPTGSGVGL-----ADGRLLR 176
 QY 176 ---AAYHPQOFIYAGLSALLD--PSQMGPSLIGLAWG-----DAGGYKAADWGPSS 224
 Db 177 LSWAATTPQFVYAGLSALLD--PSQMGPSLIGLAWG-----DAGGYKAADWGPSS 224
 QY 225 DPANERNPTQI--PKLVANNTRLVWYCGNGTNPGLGAGNIPAFLENFVRSSNLKFD 282
 Db 237 -----TRCRSASLVANNTRLVWYCGNGTNPGLGAGNIPAFLENFVRSSNLKFD 282
 QY 283 AYNAAGGHNAVFPNPGTHSWEYGAQLNAKMGDLOSLGAG 324
 Db 288 AYNAAGGHNAVFNFDANGTHDPYWGALQKAMKPDLOSLVGA 329

RESULT 5
 P71483 PRELIMINARY; PRT; 136 AA.
 ID P71483
 AC P71483
 DT 01-FEB-1997 (TremBLrel. 02, Created)
 DT 01-FEB-1997 (TremBLrel. 02, Last sequence update)
 DE 01-DEC-2001 (TremBLrel. 19, Last annotation update)
 OS 32-Kda protein (fragment).
 OS Mycobacterium asiaticum.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1790;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 25276;
 RX MEDLINE=97193847; PubMed=9041432;
 RA Soini H.K., Viljanen M.K.;
 RT "Diversity of the 32-kDa protein gene may form a basis for species
 determination of potentially pathogenic mycobacterial species."
 RL J. Clin. Microbiol. 35:769-773 (1997).
 DR EMBL; Y07715; CAA68987.1; --
 DR HSSP; P31953; 1DQZ.
 DR InterPro; IPR000801; Esterase_put.
 DR Pfam; PF00756; Esterase; 1.
 FT NON_TER 1
 FT NON_TER 136
 SQ SEQUENCE 136 AA; 14504 MW; DF1071EA4B700B42 CRC64;

Query Match 36.3%; Score 634; DB 2; Length 136;
 Best Local Similarity 83.1%; Pred. No. 8.4e-41;
 Matches 113; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 119 YSDWYSPACGKAGCQTYKWTFTLSELPWLSANRAVKPTGSAALGLSMAGSSAMILAA 178
 Db 1 YSDWYNAGCGKAGWTYKWTFTLSELPWLSANRAVKPTGSAALGLSMAGSSAMILAA 60
 QY 179 HPQOFIYAGLSALLDPSQMGPSLIGLAWDAGGYKAADWGPSSDPANERNPTQIIP 238
 Db 61 HPQOFIYAGLSALLDPSQMGPSLIGLAWDAGGYKAADWGPSSDPANERNPTQIIP 120
 QY 239 KLVANNTRLVWYCGNG 254
 Db 121 KLVANNTRLVWYCGNG 136

RESULT 6
 Q9F7W5 PRELIMINARY; PRT; 139 AA.
 ID Q9F7W5
 AC Q9F7W5
 DT 01-MAR-2001 (TremBLrel. 16, Created)
 DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
 DE 01-JUN-2002 (TremBLrel. 21, Last annotation update)
 DE Putative 32 kDa surface antigen (Fragment).
 GN FBPA.
 OS Mycobacterium ulcerans.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1809;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC19423;
 RX MEDLINE=20507800; PubMed=11053375;
 RA Steiner T.P., Jenkin G.A., Johnson P.D., Davies J.K.;
 RT "Comparative genetic analysis of mycobacterium ulcerans and
 mycobacterium marinum reveals evidence of recent divergence."
 RL J. Bacteriol. 182:6322-6330 (2000).
 DR EMBL; AF271092; AAG31743.1; --
 DR HSSP; P31953; 1DQZ.
 DR InterPro; IPR000801; Esterase_put.
 DR Pfam; PF00756; Esterase; 1.
 FT NON_TER 1
 FT NON_TER 139
 SQ SEQUENCE 139 AA; 14431 MW; 68A4138016622CF7 CRC64;

Query Match 35.6%; Score 621; DB 2; Length 139;
 Best Local Similarity 79.9%; Pred. No. 8.4e-40;
 Matches 111; Conservative 15; Mismatches 13; Indels 0; Gaps 0;

QY 157 PTGSAALGLSMAGSSAMILAAHPQOFIYAGLSALLDPSQMGPSLIGLAWDAGGYKA 216
 Db 1 PTGSGVGLSMAGSSAMILAAHPQOFIYAGLSALLDPSQMGPSLIGLAWDAGGYKA 60
 QY 217 ADMWGPSSDPANERNPTQIIPKLVANNTRLVWYCGNGTNPGLGAGNIPAFLENFVRSS 276

RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 14474;
RX MEDLINE=97193847; PubMed=9041432;
RA Soini H.K., Viljanen M.K.;
RT "Diversity of the 32-kDa protein gene may form a basis for species
RT determination of potentially pathogenic mycobacterial species.";
PI J. Clin. Microbiol. 35:769-773(1997).

Query Match	27.1%	Score 473	DB 16	Length 669
Best Local Similarity	34.9%	Pred. No. 1.2e-27		
Matches 114; Conservative	49	Mismatches 124	Indels 40	Gaps 10
QY	20	AAAVVLPLVLGLAGCAATAGAFSPGLPEVEY-----LQVPSPSM-GRDIKVV-----	65	
Db	88	AGEVTYPAIEGLPAGVRVNSA-----EYVTHSHVVLSTQSAAMPERIKVQLLLPRD	139	
QY	66	-FQSGNNSPAVYLLDGLRAODYNGNDINTAPAEFWYQSGLSITVMVGVGSSSPSYDWS	124	
Db	140	WYSSPDRDPFRIWALDGLRALEKOSGMTIETNFQFFADKNAIVLVLPGVGESSFYIDWNE	199	
QY	125	PACGKAGCQTVKWEFTLTSETLPQMLSANRAVVKPTGSAAI-GLSMAGSASAMILAAHYHPQOF	183	
Db	200	FNNGK---NQWETFTLTEELAPIL--DKGFRSNGERAITGISMGGTAAVNAIVHNPDMF	253	
QY	184	IVAGLSLALLDPSQGMGSLTLGLAMDAGGYKAADMGPSSDPDAWERNDPTQIQPKLVAN	243	
Db	254	NFVGFSFGYLDPTTSGMPEAAICAAALADAGGVNNVAMWGPAGSERWLENDPRNVQDL--R	311	
QY	244	NTRLWYVCNGTPTN-----ELGCANIPABFLENFVRSSNLKFDQAYNAAGGHNAVNF	296	
Db	312	TCGVNYSAGSDYDQCGSVATGPAANAAGVGLLEISMTTSQTFVDAANGVGNVIANF	370	

```

QY 297 PPNCTHSHWYCAQLNANKGDLQSSLG 323
DB 371 RPSGVHAWPYQFQENTQAWPYMADSLG 397

RESULT 10
P72063 PRELIMINARY; PRT; 120 AA.
AC P72063;
DT 01-FEB-1997 (TrEMBLrel. 02, Created)
DT 01-DEC-2001 (TrEMBLrel. 02, Last sequence update)
DT 01-FEB-2001 (TrEMBLrel. 19, Last annotation update)
DE 32-kDa protein (Fragment).
OS Mycobacterium vaccae.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1810;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 29678;
RX MEDLINE=97193847; PubMed=9041432;
RA Soini H.K., Viljanen M.K.;
RT "Diversity of the 32-kDa protein gene may form a basis for species
RT determination of potentially pathogenic mycobacterial species.";
RL J. Clin. Microbiol. 35:769-773 (1997).
DR EMBL; Y07719; CAA68991.1; -
DR HSSP; P31953; 1DQZ.
DR InterPro; IPR000801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
FT NON TER 1 120
FT NON TER 120 120
SQ SEQUENCE 120 AA; 13279 MW; BA935A7A28C36374 CRC64;

Query Match 24.6%; Score 428.5; DB 2; Length 120;
Best Local Similarity 65.0%; Pred. No. 3e-25;
Matches 78; Conservative 16; Mismatches 25; Indels 1; Gaps 1;

QY 136 KWETFTLSELPQWLSANRAVPTGSAIGLSMAGSSAMILAAHYHPQOFIYAGSLSLALDP 195
DB 1 KWETFTLQELFYLYQXNKGVDNENAGVLSMAGSSPVTLLAIYHPQOFQYAGSLGYLNL 60

QY 196 SQMGPSLLIGLMDAGGYKAADMWGPS-SDPAMERNPTQOIIPKLVANNTRLVWYCGNG 254
DB 61 SEGFWPMLVMSMGDAGGYDANDMWGPAETDPAWQYNDPMVQIPRLVANNRVRVWYCGNG 120

RESULT 11
Q8NLR1 PRELIMINARY; PRT; 341 AA.
AC Q8NLR1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Predicted esterase.
GN CGL2878.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005283; BAC00272.1; -
DR InterPro; IPR000801; Esterase.
DR InterPro; IPR001680; WD40.
DR Pfam; PF00756; Esterase; 1.
DR PROSITE; PS00678; WD_REPEATS_1; 1.
KW Complete proteome.
SQ SEQUENCE 341 AA; 37187 MW; 91877B75D835BE94 CRC64;

```

```

Query Match 23.8%; Score 415; DB 16; Length 341;
Best Local Similarity 33.7%; Pred. No. 1.3e-23;
Matches 116; Conservative 52; Mismatches 142; Indels 34; Gaps 9;

QY 3 DVSRKIRAWGRRLMIGTAAAVLPLGLVGIAGATAGAFSRP-----GLP 47
DB 9 EASRKLVA-----LVVALATAAAL-----MVVGQGTAAQANRDWLRADNSGYCDWDVAGFW 59

QY 48 VEYLQVPSMGRDIKQVQSGGNNSPA-VLLDGLRAQDDYNGWDINTPAFEWYQSGSL 106
DB 60 VQRCDDVSPAMGRNIPVQIQPAGRGAGLGLLDGMGRATEYSNAWLVDTNAAARYAPNNI 119

QY 107 SIVMPVGQSSFFSYDWSYSPA-CGKAGCQTYKWETFTLSELPQWLSANRAVPTGSAATGL 165
DB 120 TLVMPVGAGSFYADWNSQASLSSDPVIYWEFTLQELPAYLEQNGFVARNNNSIGGL 179

QY 166 SMAGSSAMILAAHYHPQOFIYAGSLSLALDPSQGMGPSILGLMDAGGYKAADMWGPS 225
DB 180 SMGGTAALNLAAKHPDQFQAMSWSGYLNTTAPGMQTLRLMLDLDTGCFNVNMYGSIIN 239

QY 226 PAWERNPTQOIIPKLVANNTRLVWYCGNG--TPNELG---GANIPAELENFVRSSNLKP 280
DB 240 PRRENDFPWNMGGLA--NTDVLISAASGLWSPDDGVVRVDRHTGVSYLE-FVAMTSTRI 296

QY 281 QDAYNAAAGHNAVFNPFPNGTHSWYCAQLNANKGDLQSSLG 324
DB 297 WEAKARLQGLNPTADYPMYGLHGWAFNSQLERLQGRVLDVWNA 340

RESULT 12
Q8FLZ9 PRELIMINARY; PRT; 360 AA.
AC Q8FLZ9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative P51 protein.
GN CR2710.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kwarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005223; BAC19520.1; -
KW Complete proteome.
SQ SEQUENCE 360 AA; 39009 MW; A2AFF89D8D7C4619 CRC64;

Query Match 23.2%; Score 404; DB 16; Length 360;
Best Local Similarity 34.1%; Pred. No. 9.4e-23;
Matches 118; Conservative 51; Mismatches 137; Indels 40; Gaps 10;

QY 4 VSRKIRAWGRRLMIGTAAAVLPLGLVGIAGATAG--AFSRP-----GLPVEY 50
DB 29 VSHKLLA-----LVVALATAA-----GLMVVGTTGAQAANRDWLRPDSTGHCDWDGVYVQR 81

QY 51 LQVPSMGRDIKQVQSGGNNSPA-VLLDGLRAQDDYNGWDINTPAFEWYQSGSL 107
DB 82 CDVYSPAMGRNIAVQIQPAGRGAGLGLLDGMGRATEYSNAWLVDTNAAARYAPNNI 139

QY 108 IVMVPVGQSSFFSYDWSYSPA-CGKAGCQTYKWETFTLSELPQWLSANRAVPTGSAATGL 166
DB 140 LVMPVGAGSFYADWNNHPTLSSAEFVYVWETFTLQELPAYLEQNGFVARNNNSVAGLS 199

QY 167 MAGSSAMILAAHYHPQOFIYAGSLSLALDPSQGMGPSILGLMDAGGYKAADMWGPS 226
DB 167 MAGSSAMILAAHYHPQOFIYAGSLSLALDPSQGMGPSILGLMDAGGYKAADMWGPS 226

```

```
Db 200 MGGTAALNLAAKHPQCFQAMSYSGYLTTTAPGCMQTLMLAAMDGTGGFNVMAMVGSVINP 259
QY 227 AWERNDPTQOIPKL-----VANTRLWVYCGNGT--PNELGGANIPAEFLNFVRSSNL 278
Db 260 RRFENDPFNNWGLRGKDVVYSAASGLWGPQDNGTRVDRHNGSVLEA-----VSLATT 313
QY 279 KFQDAYNAAGHNVAFFNPPNGTHSWYWGQALNMGKGLDLOSSIGA 324
Db 314 RAWAKARAEGLNVTADYPTNGIHSWAQFSSQLHKTRDRVLNVNNA 359

RESULT 13
P71536
ID P71536 PRELIMINARY; PRT; 96 AA.
AC P71536;
DT 01-FEB-1997 (TRENBLrel. 02, Created)
DT 01-FEB-1997 (TRENBLrel. 02, Last sequence update)
DT 01-JUN-2002 (TRENBLrel. 21, Last annotation update)
DE 32-kDa protein (Fragment).
OS Mycobacterium triviale.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1798;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23291; PubMed=9041432;
EX MEDLINE=97193847;
RA Soini H.K., Viljanen M.K.;
RT "Diversity of the 32-kDa protein gene may form a basis for species
determination of potentially pathogenic mycobacterial species.";
RL J. Clin. Microbiol. 35:769-773(1997).
DR EMBL: Y07716; CAA68988.1; -.
DR HSP; P31953; 1DQZ.
DR InterPro; IPR000801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
FT NON_TER 1
FT NON_TER 96
SQ SEQUENCE 96 AA; 10206 MW; 52FB7E4708FBA85E CRC64;

Query Match 19.8%; Score 346; DB 2; Length 96;
Best Local Similarity 61.5%; Pred. No. 4.2e-19;
Matches 59; Conservative 16; Mismatches 21; Indels 0; Gaps 0;

QY 159 GSAALGSMAGSSAMILAAHYHPQFIYAGSLALLDPSQGMGSPSLIGLAMGDAGYKAAD 218
Db 1 GNAVVLGSMGSGTATYALWPEQFAFAGSLSGFLNPSNGWPTLIGFAMKDGAFNGTD 60

QY 219 MWGSSDPAWERNPTQOIPKLVANTRLWVYCGNG 254
Db 61 MWGPAADVAWRNDPTVANVARLVANTRVWVYCGNG 96

RESULT 14
Q8NRL7
ID Q8NRL7 PRELIMINARY; PRT; 411 AA.
AC Q8NRL7;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein Cg11031.
GN CGL1031.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AP005277; BAB98424.1; -.
DR InterPro; IPR000801; Esterase_put.
```

```
DR InterPro; IPR000379; Ser_estr_site.
DR Pfam; PF00756; Esterase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 411 AA; 43660 MW; B9ABBC09B501FE3 CRC64;

Query Match 17.4%; Score 303.5; DB 16; Length 411;
Best Local Similarity 30.2%; Pred. No. 4.9e-15;
Matches 86; Conservative 49; Mismatches 121; Indels 29; Gaps 10;

QY 45 GLPVEYLQVPSPMGRDIKVQ-FQSGGNNSPA--VYLLDGLRAQDDYNGWDINTPAFEW 101
Db 126 GLRLEKSVASPSMQRNVDVQIMKSAEADSPAPMLYMLDGGKNNKSSGNGGEGPKVF 185
QY 102 YQSGLSIVMPVGGOSFYSDWY--SPACGKAGCQTYKWTETFTLTSELPOWLSANRAVKPTG 159
Db 186 ADENVTVVMPGLCAASSMYSDWLEEDPALGR-----IKWETIVBELAPLLAEAEELNENG 240
QY 160 SAAI-GLSMAGSSAMILAAHYHPQF---IYAGSLALLDPSQGMGSPSLIGLAMGDAGYK 215
Db 241 HRGIGLSMGATGAVHLANSNPDLFDGVIGISGCVSTLDP---IGQTTVSLIVNSRGG-N 296
QY 216 AADMWGSSDPAWERNPTQOIPKLVANTRLWVYCGNGTNELGGANIPAE----- 267
Db 297 VENMWGPTGSETWKAHDVTSNPEGL--RDMAYVLSAANGVVDLIDLADSEKEPFYNLLAG 354
QY 268 -FLENFVRSNLKFQDAYNAAGHNVAFFNPPNGTHSWYWGQAL 311
Db 355 VLERGSLSCTEALDESNRAGMHQVVDYKDSGTHNWNPNQL 399

RESULT 15
Q9R667
ID Q9R667 PRELIMINARY; PRT; 77 AA.
AC Q9R667;
DT 01-MAY-2000 (TRENBLrel. 13, Created)
DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)
DE Antigen 85B=SECRETED protein (Fragments).
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE.
EX MEDLINE=92366936; PubMed=1502498;
RA Wiker H.G., Nagai S., Harboe M., Ljungqvist L.;
RT "A family of cross-reacting proteins secreted by Mycobacterium
tuberculosis.";
RL Scand. J. Immunol. 36:307-319(1992).
DR HSP; P31953; 1DQY.
FT NON_TER 1
FT NON_TER 41
FT NON_TER 42
FT NON_TER 77
SQ SEQUENCE 77 AA; 8385 MW; C1802000ED3B1515 CRC64;

Query Match 17.0%; Score 297.5; DB 2; Length 77;
Best Local Similarity 28.8%; Pred. No. 1.5e-15;
Matches 75; Conservative 0; Mismatches 2; Indels 183; Gaps 1;

QY 41 FSRPGLPVEYLQVPSPMGRDIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEW 100
Db 1 FSRPGLPVEYLQVPSPMGRDIKVQFQSGGNNSPAVYLLDG----- 41
QY 101 YQSGLSIVMPVGGOSFYSDWYSPACGKAGCQTYKWTETFTLTSELPOWLSANRAVKPTGS 160
Db 42 ----- 41
QY 161 AAIGLSMAGSSAMILAAHYHPQFIYAGSLALLDPSQGMGSPSLIGLAMGDAGYKAADMW 220
Db 42 ----- 41
QY 221 GPSSDPAWERNPTQOIPKLVANTRLWVYCGNGTNELGGANIPAEFLNFVRSSNLKF 280
DR -----
```



```
Db 42 -----PAEFLNFVRSSNLKP 57
QY 281 QDAYNAGGHNVAFFPPNG 300
Db 58 QDAYKPAGGHNVAFFPPNG 77

RESULT 16
Q8FTS3
ID Q8FTS3 PRELIMINARY; PRT; 390 AA.
AC Q8FTS3;
DT 01-OCT-2002 (TrEMBLrel. 23, Created)
DT 01-VAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-VAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative PSI protein.
GN CE1488.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005219; BAC18298.1; -.
KW Complete proteome.
SQ SEQUENCE 390 AA; 41356 MW; 46C3B7B47B08DBF8 CRC64;

Query Match 16.3%; Score 284.5; DB 16; Length 390;
Best Local Similarity 28.6%; Pred. No. 1.3e-13;
Matches 83; Conservative 46; Mismatches 122; Indels 39; Gaps 10;

QY 45 GLPVEYLQVPSMGRIKVOFO---SGGNSPAVLLDGLRAQDDYNGWDINTPAFEVY 101
Db 105 GLRLERWTVASPSMQRNVDVQIMRSVDAGAPAPMLYLDIGIGNRNSSGWINHGQGPKEV 164
QY 102 YQSGLSIVMPVGGQSFYSWDY--SPACGKAGCQTYKWTFTLSELTPQLMSANRAVKPTG 159
Db 165 GENVIVVMPGLGAASMYSDWEEDPALGR-----IWWETFIIVEELAPLLEAEELFNFG 219
QY 160 SAAI-GLSMAGSSAMILAAVHPQOF---IYAGLSALLDPSQGMGSLIGLAMGDAGYKAAD 215
Db 220 HRGIGLSMGATGAVHLANANPDFFDAVIGISCYSTLDP---IGQATVSLIVKSRG-D 275
QY 216 AADMWGPSSDPAPERNDPQQIPKLVAN-----NTRLWYCGNGTNPENLGGANIPAEFLE 270
Db 276 VENMWGPVGSRTWQEHQ-----VVSNPEGLNMAVYLSAANGVVDEIDREYADEPFY 328
QY 271 NFVRSNLK-----FQDAYNAAGGHNVAFFPPNGTHSWEYGAOL 311
Db 329 NLLAGTVLBERGALSCTEALDDAQDAGMTHQVVDYKAGAHNRNFNEQL 378

RESULT 17
Q8NRW9
ID Q8NRW9 PRELIMINARY; PRT; 462 AA.
AC Q8NRW9;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Hypothetical protein Cg10922.
GN Cg10922.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1716;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;

Query Match 15.9%; Score 277; DB 16; Length 483;
Best Local Similarity 28.9%; Pred. No. 6.3e-13;
Matches 82; Conservative 43; Mismatches 133; Indels 26; Gaps 9;

QY 48 VEYLQVPSMGRIKVOFQSGGNS---PAYVLLDGLRAQDDYNGWDINTPAFEVY 104
Db 129 IRQINAYSPSMGRITPLVWVPEDNTPGTVYALGGDGGQGNWVTRTDLDELTSN 188
QY 105 GLSIVMPVGGQSFYSWDYSPACGKAGCQTYKWTFTLSELTPQLMSANRAVKPTGSAAI- 163
Db 189 NINLIMPMLGGSFYADWAGESSEMGGAQ--QWETFLMHELPELEA--AIGADGORSIV 244
QY 164 GLSMAGSSAMILAAVHPQOFIYAGLSALLDPSQGMGSLIGLAMGDAGYKAADMWGPS 223
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;

Query Match 16.0%; Score 279; DB 16; Length 462;
Best Local Similarity 28.9%; Pred. No. 4.2e-13;
Matches 82; Conservative 44; Mismatches 132; Indels 26; Gaps 9;

QY 48 VEYLQVPSMGRIKVOFQSGGNS---PAYVLLDGLRAQDDYNGWDINTPAFEVY 104
Db 108 IRQINAYSPSMGRITPLVWVPEDNTPGTVYALGGDGGQGNWVTRTDLDELTSN 167
QY 105 GLSIVMPVGGQSFYSWDYSPACGKAGCQTYKWTFTLSELTPQLMSANRAVKPTGSAAI- 163
Db 168 NINLIMPMLGGSFYADWAGESSEMGGAQ--QWETFLMHELPELEA--AIGADGORSIV 223
QY 164 GLSMAGSSAMILAAVHPQOFIYAGLSALLDPSQGMGSLIGLAMGDAGYKAADMWGPS 223
Db 224 GMSMGSGSVINPETHDPNIFYSSVGSFGCAETNSWMGRR--GIATAYNGNVVPEQIFGE 281
QY 224 SDPAWER-NDPTQQIPKLVANNTLWVYCGNGTNPENLG--GANIPAE-----F 268
Db 282 VDSYDSRYNDPLNNAKL-EEQDNLVIFAGSGVSELDVIGNAPIDEDEAFKRVLVGFE 340
QY 269 LENFVRSNLKFDQDAYNAAGGHNVAFFPPNGTHSWEYGAOLN 312
Db 341 IEAMSNCTCTHNLKAATDQMGIDNINIDFRPTCTHAWDYWNEALH 384

RESULT 18
Q8NNL2
ID Q8NNL2 PRELIMINARY; PRT; 483 AA.
AC Q8NNL2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein Cg12181.
GN Cg12181.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005280; BAB99574.1; -.
DR InterPro; IPR000801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 483 AA; 51567 MW; 8449C06238D83230 CRC64;

Query Match 15.9%; Score 277; DB 16; Length 483;
Best Local Similarity 28.9%; Pred. No. 6.3e-13;
Matches 82; Conservative 43; Mismatches 133; Indels 26; Gaps 9;

QY 48 VEYLQVPSMGRIKVOFQSGGNS---PAYVLLDGLRAQDDYNGWDINTPAFEVY 104
Db 129 IRQINAYSPSMGRITPLVWVPEDNTPGTVYALGGDGGQGNWVTRTDLDELTSN 188
QY 105 GLSIVMPVGGQSFYSWDYSPACGKAGCQTYKWTFTLSELTPQLMSANRAVKPTGSAAI- 163
Db 189 NINLIMPMLGGSFYADWAGESSEMGGAQ--QWETFLMHELPELEA--AIGADGORSIV 244
QY 164 GLSMAGSSAMILAAVHPQOFIYAGLSALLDPSQGMGSLIGLAMGDAGYKAADMWGPS 223
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
```

```
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005276; BAB98315.1; -.
DR InterPro; IPR000801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 462 AA; 49365 MW; 4CC76BCE6A8C4B9E CRC64;

Query Match 16.0%; Score 279; DB 16; Length 462;
Best Local Similarity 28.9%; Pred. No. 4.2e-13;
Matches 82; Conservative 44; Mismatches 132; Indels 26; Gaps 9;

QY 48 VEYLQVPSMGRIKVOFQSGGNS---PAYVLLDGLRAQDDYNGWDINTPAFEVY 104
Db 108 IRQINAYSPSMGRITPLVWVPEDNTPGTVYALGGDGGQGNWVTRTDLDELTSN 167
QY 105 GLSIVMPVGGQSFYSWDYSPACGKAGCQTYKWTFTLSELTPQLMSANRAVKPTGSAAI- 163
Db 168 NINLIMPMLGGSFYADWAGESSEMGGAQ--QWETFLMHELPELEA--AIGADGORSIV 223
QY 164 GLSMAGSSAMILAAVHPQOFIYAGLSALLDPSQGMGSLIGLAMGDAGYKAADMWGPS 223
Db 224 GMSMGSGSVINPETHDPNIFYSSVGSFGCAETNSWMGRR--GIATAYNGNVVPEQIFGE 281
QY 224 SDPAWER-NDPTQQIPKLVANNTLWVYCGNGTNPENLG--GANIPAE-----F 268
Db 282 VDSYDSRYNDPLNNAKL-EEQDNLVIFAGSGVSELDVIGNAPIDEDEAFKRVLVGFE 340
QY 269 LENFVRSNLKFDQDAYNAAGGHNVAFFPPNGTHSWEYGAOLN 312
Db 341 IEAMSNCTCTHNLKAATDQMGIDNINIDFRPTCTHAWDYWNEALH 384

RESULT 18
Q8NNL2
ID Q8NNL2 PRELIMINARY; PRT; 483 AA.
AC Q8NNL2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein Cg12181.
GN Cg12181.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005280; BAB99574.1; -.
DR InterPro; IPR000801; Esterase_put.
DR Pfam; PF00756; Esterase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 483 AA; 51567 MW; 8449C06238D83230 CRC64;

Query Match 15.9%; Score 277; DB 16; Length 483;
Best Local Similarity 28.9%; Pred. No. 6.3e-13;
Matches 82; Conservative 43; Mismatches 133; Indels 26; Gaps 9;

QY 48 VEYLQVPSMGRIKVOFQSGGNS---PAYVLLDGLRAQDDYNGWDINTPAFEVY 104
Db 129 IRQINAYSPSMGRITPLVWVPEDNTPGTVYALGGDGGQGNWVTRTDLDELTSN 188
QY 105 GLSIVMPVGGQSFYSWDYSPACGKAGCQTYKWTFTLSELTPQLMSANRAVKPTGSAAI- 163
Db 189 NINLIMPMLGGSFYADWAGESSEMGGAQ--QWETFLMHELPELEA--AIGADGORSIV 244
QY 164 GLSMAGSSAMILAAVHPQOFIYAGLSALLDPSQGMGSLIGLAMGDAGYKAADMWGPS 223
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
```

Db 245 GMSGSGVLNFAHDFNFYSSVSGCAETNSWMCRR--GTAATAYNGNVVPEQIFGE 302

QY 224 SDPAWER-NDEPTQIPLKLVANNTELWYCGTGNELG--GANIPAE-----F 268

Db 303 VDSYRYNDPLNNAKL-EQDNLXIFAGSGVSELDVIGDNAPIDEDAFKKNVLVQFE 361

QY 269 LENFVRSNKLKFDAYNAAGHNAVFPPNPGNTHSWYGAQLN 312

Db 362 IEAMSNTCHNLKAAATQMGIDNINIDFRPTGTHAWDYNEALH 405

RESULT 19

Q93R90 PRELIMINARY; PRT; 358 AA.

AC Q93R90;

DT 01-DEC-2001 (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)

DE Surface layer protein A.

GN SLPA.

OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.

OX NCBI_TaxID=1697;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 6872;

RA Usuda Y., Kawasaki H., Utagawa T.;

RT "Characterization of a surface layer protein of Corynebacterium

RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AB05224; BAB62413.1; -

DR InterPro; IPR000801; Esterase_cat.

DR Pfam; PF00756; Esterase; 1.

FT CHAIN 26 358

FT MATURE SURFACE LAYER PROTEIN A.

SQ SEQUENCE 358 AA; 39081 MW; 0785CA6602351A1C CRC64;

Query Match 15.1%; Score 263; DB 2; Length 358;

Best Local Similarity 27.5%; Pred. No. 4.9e-12;

Matches 81; Conservative 46; Mismatches 130; Indels 38; Gaps 10;

QY 51 LQVPSM-GRDIKVFQSGN-NSPAVLLDGLRAQDYNWDINTPAFVYQSLGI 108

Db 69 LQATSPAMDKRVLAIIRAQNPDRPTIYLLNGAGSBDTDLNQSEAVDFYADKQNV 128

QY 109 VMPVGGSSFFSDVSPACGAGCQTYKWFELTSELPOWLSANRAVPTGSAALGMA 168

Db 129 VIPQAGAFSYTDNTTTPNKSYLKGPKQWETFLTKELPGPLE-ERLOSNNKRAIGMSMS 187

QY 169 GSSAMILAAVHPQOFIVAGSLISALLDPSQGMGSLIGLAMGDAGGYKAADNWGSPSDFAW 228

Db 188 ATSSILLAHQNGFYDVGVSAGCAGTSTPEYEAMRLTVNRGG-EPQMWKMGSETN 246

QY 229 ERNPTQIPLKLVANNTELWYCGTGNELGGANIPAEFLN-----FVRSNKLKFQ 281

Db 247 RYNDALLNSDKL---RGTYLYSSGNLPGE---TDMPSYTKQGVDPDTTASVGAATLQIE 301

QY 282 DAYNAG---GHN-----AVENFPNCTHSEWYGAQLNAMKGLQSS 321

Db 302 GGIETAGVNHCTHLEAKLKSQNTPAIYFRDGTGTHSWPGW-----REDLEKS 349

RESULT 20

Q9S6X0 PRELIMINARY; PRT; 52 AA.

AC Q9S6X0;

DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Alpha antigen, antigen 85B, MPB59, antigen 6 precursor (Fragment).

OS Corynebacterium bovis.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1765;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=BCG Tokyo;

RX MEDLINE=88314872; PubMed=2842287;

RA Matsuo K., Yamaguchi R., Yamazaki A., Tasaka H., Yamada T.;

RT "Cloning and expression of the Mycobacterium bovis BCG gene for

RT extracellular alpha antigen.";

RL J. Bacteriol. 170:3847-3854(1988).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=BCG Tokyo;

RX MEDLINE=98070721; PubMed=9405208;

RA Ohara N., Nishiyama T., Ohara-Wada N., Matsumoto S., Matsuo T.;

RA Yamada T.;

RT "Characterization of the transcriptional initiation regions of genes

RT for the major secreted protein antigens 85C and MPB51 of Mycobacterium

RT bovis BCG.";

RL Microb. Pathog. 23:303-310(1997).

DR EMBL; D78142; BAA11219.1; -

KT Signal.

FT SIGNAL 1 40 POTENTIAL.

FT CHAIN 41 >52 ALPHA ANTIGEN, ANTIGEN 85B, MPB59,

FT ANTIGEN 6.

FT NON TER 52 52

SQ SEQUENCE 52 AA; 5324 MW; 4DB001D1DF9CB082 CRC64;

Query Match 14.9%; Score 260; DB 2; Length 52;

Best Local Similarity 100.0%; Pred. No. 6.5e-13;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGCAATAGAFSRPGLPVEYLQ 52

Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGCAATAGAFSRPGLPVEYLQ 52

RESULT 21

Q8FSM7 PRELIMINARY; PRT; 381 AA.

AC Q8FSM7;

DT 01-MAR-2003 (TrEMBLrel. 23, Created)

DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE Conserved hypothetical protein.

GN C80356.

OS Corynebacterium efficiens.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.

OX NCBI_TaxID=152794;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;

RA Kawarayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.;

RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.;

RA Usuda Y., Sugimoto S.;

RT "The entire genomic sequence of Corynebacterium efficiens YS-314.";

RT Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AP005215; BAC17186.1; -

KT Hypothetical protein; Complete proteome.

SW Hypothetical protein; Complete proteome.

SQ SEQUENCE 381 AA; 41563 MW; 5629AF5EFDCE79BE CRC64;

Query Match 14.8%; Score 257.5; DB 16; Length 381;

Best Local Similarity 25.5%; Pred. No. 1.4e-11;

Matches 93; Conservative 45; Mismatches 166; Indels 61; Gaps 13;

QY 13 RRLMIGTAA-----AVVLEPLGLVLA--GGAATAGAFSRPGLPV--EYLQVP----- 54

Db 18 RRIAAPVAALGIALSTVVAPTAVSAALTPAAVAGDTQSTITVCAKYEQPAPRWARV 77

QY 55 -----SPSMGRDIK---VQFQSGNNSPAVYLLDGLAQDYNWDINTPAF 98

Db 78 NASGERVKEMWAYSPSMDRDPVLVITADESAGRPVYLLNGDGGEGGNANWIMQTDVT 137

```

QY 99 EMYOSGLSIVMPVGGSSFYDWYSPACGKAGCQTYKWTETLSELPOWLSANRAVKPT 158
Db 138 DYLEKNNVNLPMGKFSYTDWQENALGGKQ--MWETFLVKELPGPL--EBELNAD 193
QY 159 GSAAT-GLSMAGSSAMILAAHPOOFIYAGLSALLDPSQGMPSLIGLWAGDAGYKAA 217
Db 194 GQRAIAGMSMATSLLPQHPYGFYDAAASFGSCASTSQPLPWEYIRLTL--DRGNATPE 252
QY 218 DMWGSPPAWERNPTQOIPKLVANTRLWYCGNTPNELGGANIP-----AEFLE- 270
Db 253 QMWGPRGGEVNIYNDALNSDKL--RGTDLYISNAGSLAGHESANSFENGLDQAYLSL 310
QY 271 -----NFRSSNLKFQDAYNAAGHNAV---FNEPPNGTHSWEYWGQALNMGKDL 318
Db 311 AMTETIVTGGLEIATNCTHDLKAKLDHAGIPADWNLRTPTGTHSWGWDLRGSDWTF 370
QY 319 QSSLG 323
Db 371 ARSFG 375

RESULT 22
Q8FQY3
ID Q8FQY3 PRELIMINARY; PRT; 484 AA.
AC Q8FQY3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Putative surface layer protein.
GN CE0984.
OS Corynebacterium efficiens.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=152794;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;
RA Kawarabayasi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,
RA Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,
RA Usuda Y., Sugimoto S.;
RT "The entire genomic sequence of Corynebacterium efficiens YS-314."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005217; BAC17794.1;
KW Complete proteome.
SQ SEQUENCE 484 AA; 51867 MW; 5DF15F5A933478EB CRC64;

Query Match 14.4%; Score 252; DB 16; Length 484;
Best Local Similarity 28.4%; Pred. No. 5.1e-11;
Matches 82; Conservative 42; Mismatches 135; Indels 30; Gaps 10;

QY 45 GLPVEYLQVPSMGRDIKVF---QSGGNNSPAVYLLDGLRAQDDYNGWDINTPAPEWY 101
Db 122 GERIQINAYSPMERWIPLVWIVPDTSEPRTLVYALGGDQGSANNITKTDPELM 181
QY 102 YOSGLSTVMPVGGSSFYDWYSPACGKAGCQTYKWTETLSELPOWLSANRAVKPTGS- 160
Db 182 SSNNHVIMPLGSHFYADWVEENDSLGGKQ--QWETFLTHPEPLEA--AIGEDGQR 237
QY 161 AATGLSMAGSSAMILAAHPOOFIYAGLSALLDPSQGMPSLIGLWAGDAGYKAAW 220
Db 238 SLIGMSGGSVNIAHQPNFVSSVASLSGCAETSWMRRGVA--ATVYSGNATPTQIF 296
QY 221 GPSDDPWERNDPTQOIPKLVANTRLWYCGNTPNELG--GANIP----- 265
Db 297 GEVDSYARYNDPVINAHRL-AQKDNLVFAASGVSEVDVEGENAPEDEKGNKITVG 355
QY 266 --AEFLENFRSSNLKFQDAYNAAGHNAVFNFPNGTHSWEYWGQALN 312
Db 356 FRIEALSN--TCTHNLKAADTYH--GIDTIHYDFRPTGTHAWDYNEALH 401

RESULT 23

```

```

Q8NTG4
ID Q8NTG4 PRELIMINARY; PRT; 365 AA.
AC Q8NTG4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Hypothetical protein Cgl0343.
GN CGL0343.
OS Corynebacterium glutamicum (Brevibacterium flavyum).
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
RA Nakagawa S.;
RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP005275; BAB97736.1;
DR Interpro; IPR000801; Esterase_1.
DR Pfam; PF00756; Esterase_1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 365 AA; 39506 MW; 57A7BF984A60D6C CRC64;

Query Match 13.7%; Score 238.5; DB 16; Length 365;
Best Local Similarity 26.9%; Pred. No. 3.7e-10;
Matches 78; Conservative 42; Mismatches 133; Indels 37; Gaps 10;

QY 55 SPSMGRIK---VOFGGNNSPAVYLLDGLRAQDDYNGWDINTPAPEWYQSGLSIVMP 111
Db 78 SPSMDRNVPLVITTADESAGRPVYILLNGDGGGGAANWVMTDVLDFYLEKNNVVIP 137
QY 112 VGGSSFYDWYSPACGKAGCQTYKWTETLSELPOWLSANRAVKPTGSAI--GLSMAGS 170
Db 138 MEGKFSYTDWVENASLGGKQ--MWETFLVKELPGPL--EELNTDQORAIAGMSKAT 193
QY 171 SAMILAAHPQOFTIYAGLSALLDPSQGMPSLIGLWAGDAGYKAAADWGPSSDPAWER 230
Db 194 TSLLPQHPFGFYDAAASFGSCATSSLLPWEYIKLTL--DRGNATPQMGPRGGEYNIY 252
QY 231 NDPTQOIPKLVANTRLWYCGNTPNELGGANIPA-EFLENFRSSNLK-----FQ 281
Db 253 NDALNSDKL--RGTELYVSNAGSLAGWESVDSRPEGLNQOVQSIAMAEVTVTGGIIE 310
QY 282 DAYNA-----AGGHNAVFNPNGTHSWEYWGQALNMGKDLQSS 321
Db 311 AATNKCTHDLKAKLDSAGIPADWNLRTPTGTHSWGWW-----QDDLRS 353

RESULT 24
Q86543
ID Q86543 PRELIMINARY; PRT; 342 AA.
AC Q86543;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Putative secreted protein.
GN SCO6609 OR SC1F2.06.
OS Streptomyces coelicolor.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]

```

SEQUENCE FROM N.A.
STRAIN=A3(2);
MEDLINE=97000351; PubMed=88433436;
Redenbach M., Kieser H.M., Denapaite D., Eichner A., Cullum J.,
Kinasli H., Hopwood D.A.;
"A set of ordered cosmids and a detailed genetic and physical map for
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
Mol. Microbiol. 21:77-96(1996).
[4]
SEQUENCE FROM N.A.
STRAIN=A3(2) / M145;
MEDLINE=21996410; PubMed=12000953;
Bentley S.D., Charter K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
Huang C.-H., Kieser T., Marke L., Murphy L., Oliver K., O'Neill S.,
Rabinowitch E., Rajandream M.A., Rutherford K., Rutter S.,
Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
Hopwood D.A.;
"Complete genome sequence of the model actinomycete Streptomyces
coelicolor A3(2).";
Nature 417:141-147(2002).
ENEL; AL939128; CAA20497.1; -;
InterPro; IPR000801; Esterase_put.
InterPro; IPR000379; Ser_estrif_site.
Pfam; PF00756; Esterase; 1.
Complete proteome.
SEQUENCE 342 AA; 37450 MW; 1FA7C5635BB21435 CRC64;
SO

[illegible]

RESULT 25	
Q9RSJ7	PRELIMINARY; PRT; 39 AA.
ID	
AC	Q9RSJ7;
DT	01-MAY-2000 (TrEMBLrel. 13, Created)
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Antigen 85A (Fragment).
OS	Mycobacterium tuberculosis.
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC	Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX	NCBI_TaxID=1773;
RN	[1]

```

RP SEQUENCE.
RX MEDLINE=92366936; PubMed=1502498;
RA Wiker H.G., Nagai S., Harboe M., Ljungqvist L.;
RW "A family of cross-reacting proteins secreted by Mycobacterium
RT tuberculosis."
RT Scand. J. Immunol. 36:307-319 (1992).
DR HSSP: P31953; IDQY..
DR InterPro: IPR00801; Esterase_put.
DR Pfam: PF00756; Esterase; 1.
SQ SEQUENCE 39 AA; 4221 MW; 6E7684361E1BE844 CRC64;

Query Match 11.0%; Score 192; DB 2; Length 39;
Best Local Similarity 94.9%; Pred. No. 6.5e-08;
Matches 37; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 41 FSRPGLPVEVYLQVPSMGRDIKVQFSQGNNSPAVYLL 79
   |||||
DB 1 FSRPGLPVEVYLQVPSMGRDIKVQFSQGNNSPAVYLL 39

RESULT 26
Q93JZ6 PRELIMINARY; PRT; 50 AA.
ID Q93JZ6
AC Q93JZ6;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (T-EMBLrel. 21, Last annotation update)
DE Ag85A (Fragment).
GN FBPA.
OS Mycobacterium intracellulare.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1767;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DO69;
RA Maria G.J.;
RW Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

```

```

RA Menendez C.; Characterization of a M. intracellulare variant strain by molecular
RT "techniques.";
RT Submitted (AUG-2001) to the EMBL/GenBank/DDBJ databases.
RL EMBL; AJ306712; CAC51392.1; -.
DR InterPro; IPR00801; Esterase_put.
DR P-am; PF00756; Esterase; 1.
DR NON TER 1
FT NON TER 50
SQ SEQUENCE 50 AA; 5170 MW; 6B6071C61B3364676 CRC64;

Query March 9.6%; Score 167; DB 2; Length 50;
Best Local Similarity 66.7%; Pred. No. 7.2e-06;
Matches 32; Conservative 7; Mismatches 9; Indels 0; Gaps 0;

Oy 169 GSSAMILAAVHPQOFIYAGSISALLDPSCQMGRSLIGIAMGDAGGYKA 216
      1 GGSLLITIAAYVPQPFYAALSGFINPSEGMWPTLIIGIAMSDSGSYSA 48

```

RESULT 27	
Q99ZY1	
ID	PRELIMINARY; PRT; 261 AA.
AC	Q99ZY1
DC	Q99ZY1, 2001 (TrEMBLrel. 17, Created)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE	Putative tributyltin esterase.
GN	ESTA OR SPY1022.
OS	Streptococcus pyogenes.
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC	Streptococcus.

```

OX NCBI_TaxID=1314;
RN
RP SEQUENCE FROM N.A.
RC STRAIN=SF370 / ATCC 700294 / Serotype M1;
RX MEDLINE=21192684; PubMed=11296296;
RA Ferreretti J.J., McShan W.M., Ajdic D.J., Savic G., Lyon K.,
RA Primeaux C., Szatze S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,
RA Qian Y., Jia H.G., Naifar F.Z., Ren Q., Zhu H., Song L., White J.,
RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663 (2001).
DR EMBL: AB006547; AA033917.1; -
DR InterPro: IPR000801; Esterase_put.
DR InterPro: IPR000379; Ser_estr_site.
DR Pfam: PF00756; Esterase; 1.
KW Complete proteome.
SQ SEQUENCE 261 AA; 30033 MW; D68C339DBB658BD CRC64;

Query Match 9.6%; Score 167; DB 16; Length 261;
Best Local Similarity 25.1%; Pred. No. 6.5e-05;
Matches 73; Conservative 39; Mismatches 109; Indels 70; Gaps 15;

QY 46 LPVEYLQVPSMGGRDIKQVF-----QSGGNNSPAVYLLDGLRAQDDYNGWDINTP 96
Db 4 IAEIHSVVL-GMERKVNVIYDQSEIPKQDQKDIPIVYLLHGMGNE--NSWQKRTA 60
QY 97 AFEWYQSGLSIVMPVGGQSFYSWSPACGKAGCQTYKWTFTLSELQWLISA---NR 153
Db 61 IERLLRHNTLIVMP-----STDLGWYTDTA--YGLNYR---ALSQELPQVLAAPFPNM 110
QY 154 AVKPTGSAAGLSMAGSSAMILAAHYHPOQFIYAGLSALLDPSQGMGPSLIGLAMDAGG 213
Db 111 TQKREKTFVAGLSMGGYGA-FKWLKSNRFSVAASFSGALDFSP---ETLLEGNLGEAY 166
QY 214 YKAADMWGPSSDPAPERNDPTQIQPKLVANN---TRLWYCGNGTPELGGANIPAEFLE 270
Db 167 WQG--VFQGFDDPLDKH---YLNKMYAESDGKTKFYAWCG-----YE 204
QY 271 NFVRSSNLKFQDAYNAAG-----GHNAVFNPNGTHSWYWGQALNAM 314
Db 205 DFLFATNEKATADFAQGLDIDYHKGH-----GKHEWYWNQQLLEV 246

RESULT 28
Q8P189 PRELIMINARY; PRT; 261 AA.
AC Q8P189;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DE Putative tributyrin esterase.
GN ESTA OR SPYM18 1005 OR SPYM3 0658.
OS Streptococcus pyogenes (serotype M18), and
OS Streptococcus pyogenes (serotype M3).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=186103, 198466;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES-S.pyogenes; STRAIN=MGAS8232 / Serotype M18;
RX MEDLINE=21927593; PubMed=11917108;
RA Smoot J.C., Barbican K.D., Van Gompel J.J., Smoot L.M., Chaussee M.S.,
RA Sylva G.L., Sturdevant D.E., Ricklefs S.M., Porcella S.F.,
RA Parkins L.D., Beres S.B., Campbell D.S., Smith T.M., Zhang Q.,
RA Kapur V., Daly J.A., Veasy L.G., Musser J.M.;
RT "Genome sequence and comparative microarray analysis of serotype M18
RT group A Streptococcus strains associated with acute rheumatic fever
RT outbreaks.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:4668-4673 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES-S.pyogenes; STRAIN=MGAS315 / Serotype M3;
RX MEDLINE=22133808; PubMed=12122206;

```

```

RA Beres S.B., Sylva G.L., Barbican K.D., Lei B., Hoff J.S.,
RA Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,
RA Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,
RA Schlievert P.M., Musser J.M.;
RT "Genome sequence of a serotype M3 strain of group A Streptococcus:
RT phage-encoded toxins, the high-virulence phenotype, and clone
RT emergence.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083 (2002).
DR EMBL: AE010028; AA0197642.1; -
DR EMBL: AE014150; AA079265.1; -
DR InterPro: IPR000801; Esterase_put.
DR InterPro: IPR000379; Ser_estr_site.
DR Pfam: PF00756; Esterase; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 261 AA; 30019 MW; D2DD768DFA628E7A CRC64;

Query Match 9.5%; Score 166; DB 16; Length 261;
Best Local Similarity 25.1%; Pred. No. 7.7e-05;
Matches 73; Conservative 39; Mismatches 109; Indels 70; Gaps 15;

QY 46 LPVEYLQVPSMGGRDIKQVF-----QSGGNNSPAVYLLDGLRAQDDYNGWDINTP 96
Db 4 IAEIHSVVL-GMERKVNVIYDQSEIPKQDQKDIPIVYLLHGMGNE--NSWQKRTA 60
QY 97 AFEWYQSGLSIVMPVGGQSFYSWSPACGKAGCQTYKWTFTLSELQWLISA---NR 153
Db 61 IERLLRHNTLIVMP-----STDLGWYTDTA--YGLNYR---ALSQELPQVLAAPFPNM 110
QY 154 AVKPTGSAAGLSMAGSSAMILAAHYHPOQFIYAGLSALLDPSQGMGPSLIGLAMDAGG 213
Db 111 TQKREKTFVAGLSMGGYGA-FKWLKSNRFSVAASFSGALDFSP---ETLLEGNLGEAY 166
QY 214 YKAADMWGPSSDPAPERNDPTQIQPKLVANN---TRLWYCGNGTPELGGANIPAEFLE 270
Db 167 WQG--VFQGFDDPLDKH---YLNKMYAESDGKTKFYAWCG-----YE 204
QY 271 NFVRSSNLKFQDAYNAAG-----GHNAVFNPNGTHSWYWGQALNAM 314
Db 205 DFLFATNEKATADFAQGLDIDYHKGH-----GKHEWYWNQQLLEV 246

RESULT 29
Q9R505 PRELIMINARY; PRT; 37 AA.
AC Q9R505;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE Antigen 85C (Fragment).
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE.
RX MEDLINE=92366936; PubMed=1502498;
RA Wiker H.G., Nagai S., Harboe M., Ljungqvist L.;
RT "A family of cross-reacting proteins secreted by Mycobacterium
RT tuberculosis.";
RL Scand. J. Immunol. 36:307-319 (1992).
DR HSSP: P31953; 1DQV.
DR InterPro: IPR000801; Esterase_put.
DR Pfam: PF00756; Esterase; 1.
SQ SEQUENCE 37 AA; 4016 MW; 08574C9E503DD953 CRC64;

Query Match 8.9%; Score 156; DB 2; Length 37;
Best Local Similarity 84.6%; Pred. No. 3.3e-05;
Matches 33; Conservative 1; Mismatches 3; Indels 2; Gaps 1;

QY 41 FSRPGLPVEYLQVPSMGGRDIKQVFQSGGNNSPAVYLL 79
Db 1 FSRPGLPVEYLQVPSMGGRDIKQVFQGGPH--AVYLL 37

```

Fri Feb 6 13:54:48 2004

```
RESULT 30
Q8VKB0 PRELIMINARY; PRT; 94 AA.
AC Q8VKB0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein MT0973.
GN MT0973.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE006982; AAK45221.1; -.
DR TIGR; MT0973; -.
KW Hypothetical protein.
SQ SEQUENCE 94 AA; 10869 MW; 32B9F4BA930DED99 CRC64;
Query Match 8.3%; Score 144.5; DB 16; Length 94;
Best Local Similarity 43.2%; Pred. No. 0.00086;
Matches 35; Conservative 3; Mismatches 26; Indels 17; Gaps 3;
QY 246 RLWVYCGNTPN-ELGGANIPAEFLENFVRSSNLKFDAYNAGGHNAVFNPNGTHSW 304
DB 11 RIWVSGNGTSPSAELSGNRRQA-----LQSKYQAA---QRRFNEFDGETHW 54
QY 305 EYWGQALNMGDLOSLGAG 325
DB 55 AYWRAPLOANLPDLQRVLGDG 75
```

Search completed: February 5, 2004, 17:40:07
Job time : 52.2024 secs

APPLICANT: WATSON, JAMES D.
APPLICANT: TAN, PAUL L.J.
TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fast-SEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,080
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1007
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-997-080-33

Query Match 100.0%; Score 1745; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Qy 61 DIKVOFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFS 120
Db 61 DIKVOFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFS 120
Qy 121 DWYSPACGKAGCQYKWEFTLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
Db 121 DWYSPACGKAGCQYKWEFTLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
Qy 181 QOFTYAGLSALLDPSQGMGSLIGLAMDGAGGYKAADMWGPSSDPPOWERNDPTQOIPKL 240
Db 181 QOFTYAGLSALLDPSQGMGSLIGLAMDGAGGYKAADMWGPSSDPPOWERNDPTQOIPKL 240
Qy 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVSSNLKFQDAYNAAGGHNAVFNPNG 300
Db 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVSSNLKFQDAYNAAGGHNAVFNPNG 300
Qy 301 THSWEYGAQINAMKGDLOSSLGAG 325
Db 301 THSWEYGAQINAMKGDLOSSLGAG 325

RESULT 2
US-08-997-362-33
; Sequence 33, Application US/08997362
; Patent No. 5985287

GENERAL INFORMATION:
APPLICANT: Tan, Paul
APPLICANT: Hiyama, Jun
APPLICANT: Visser, Elizabeth
APPLICANT: Skinner, Margot
APPLICANT: Scott, Linda
APPLICANT: Prestidge, Ross
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
NUMBER OF SEQUENCES: 194
CORRESPONDENCE ADDRESS:
ADDRESSEE: Law Offices of Ann W. Speckman
STREET: 2601 Elliott Avenue, Suite 4185
CITY: Seattle
STATE: WA
COUNTRY: USA
ZIP: 98121
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: Fast-SEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/997,362
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
FILING DATE: June 12, 1997
APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
FILING DATE: August 29, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002c2
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 33:
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-997-362-33

Query Match 100.0%; Score 1745; DB 2; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Qy 61 DIKVOFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFS 120
Db 61 DIKVOFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFS 120
Qy 121 DWYSPACGKAGCQYKWEFTLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
Db 121 DWYSPACGKAGCQYKWEFTLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
Qy 181 QOFTYAGLSALLDPSQGMGSLIGLAMDGAGGYKAADMWGPSSDPPOWERNDPTQOIPKL 240
Db 181 QOFTYAGLSALLDPSQGMGSLIGLAMDGAGGYKAADMWGPSSDPPOWERNDPTQOIPKL 240
Qy 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVSSNLKFQDAYNAAGGHNAVFNPNG 300
Db 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVSSNLKFQDAYNAAGGHNAVFNPNG 300
Qy 301 THSWEYGAQINAMKGDLOSSLGAG 325


```
Db 301 THSWEYGAQLNAMKGDLOSSLGAG 325
|||||
RESULT 3
US-08-873-970-33
; Sequence 33, Application US/08873970
; Patent No. 6001361
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,970
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002C1
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-873-970-33

Query Match 100.0%; Score 1745; DB 3; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MTDVSRKIRAWGRRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
Db 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
QY 121 DWYSPACGKAGCCTYKWEFTLSELQWL SANRAVKTGSAATGLSWAGSSAMILAAHP 180
Db 121 DWYSPACGKAGCCTYKWEFTLSELQWL SANRAVKTGSAATGLSWAGSSAMILAAHP 180
QY 181 QQFIYAGSLALLDPQSGGSPSLIGLAWGDAGGYKAADMWGSPSSDPAWERNPTQIQPKL 240
Db 181 QQFIYAGSLALLDPQSGGSPSLIGLAWGDAGGYKAADMWGSPSSDPAWERNPTQIQPKL 240

QY 241 VANNTLWVYCGNGTNPENLGGANI PAEFLFNVRSSNLKFO DAYNAAGGHNAVFPPNG 300
Db 241 VANNTLWVYCGNGTNPENLGGANI PAEFLFNVRSSNLKFO DAYNAAGGHNAVFPPNG 300
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
Db 301 THSWEYGAQLNAMKGDLOSSLGAG 325

RESULT 4
US-09-095-855-33
; Sequence 33, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TREATMENT AND DIAGNOSIS OF Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002C3
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-095-855-33

Query Match 100.0%; Score 1745; DB 3; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MTDVSRKIRAWGRRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
Db 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
```

Db 61 DIKVOFSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGQSSFFS 120
QY 121 DWYSPACGKAGCQTYKWETFLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
Db 121 DWYSPACGKAGCQTYKWETFLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
QY 181 QOFTYAGSLALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNPTQOIPKL 240
Db 181 QOFTYAGSLALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNPTQOIPKL 240
QY 241 VANNTRLWVYCGNGTNPNEGNIIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNPNG 300
Db 241 VANNTRLWVYCGNGTNPNEGNIIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNPNG 300
QY 301 THSWEYGAQNLNAMKGDLOSSLGAG 325
Db 301 THSWEYGAQNLNAMKGDLOSSLGAG 325

RESULT 6
US-09-324-542-33
; Sequence 33, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c1
; CURRENT APPLICATION NUMBER: US/09/324,542
; CURRENT FILING DATE: 1999-06-02
; EARLIER APPLICATION NUMBER: US 08/997,080
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-324-542-33

Query Match 100.0%; Score 1745; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVOFSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGQSSFFS 120
Db 61 DIKVOFSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGQSSFFS 120
QY 121 DWYSPACGKAGCQTYKWETFLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
Db 121 DWYSPACGKAGCQTYKWETFLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
QY 181 QOFTYAGSLALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNPTQOIPKL 240
Db 181 QOFTYAGSLALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNPTQOIPKL 240
QY 241 VANNTRLWVYCGNGTNPNEGNIIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNPNG 300
Db 241 VANNTRLWVYCGNGTNPNEGNIIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNPNG 300
QY 301 THSWEYGAQNLNAMKGDLOSSLGAG 325
Db 301 THSWEYGAQNLNAMKGDLOSSLGAG 325

RESULT 7
US-08-690-347-1

QY 121 DWYSPACGKAGCQTYKWETFLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
Db 121 DWYSPACGKAGCQTYKWETFLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
QY 181 QOFTYAGSLALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNPTQOIPKL 240
Db 181 QOFTYAGSLALLDPSQGMGSLIGLAMGAGGYKAADMWGPSSDPDAWERNPTQOIPKL 240
QY 241 VANNTRLWVYCGNGTNPNEGNIIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNPNG 300
Db 241 VANNTRLWVYCGNGTNPNEGNIIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNPNG 300
QY 301 THSWEYGAQNLNAMKGDLOSSLGAG 325
Db 301 THSWEYGAQNLNAMKGDLOSSLGAG 325

RESULT 5
US-08-705-347A-33
; Sequence 33, Application US/08705347A
; Patent No. 6284255
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESSES:
; ADDRESSES: Speckman Picard PLLC
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/705,347A
; FILING DATE: 28-AUG-1996
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206.269.0565
; TELEFAX: 206.269.0563
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-705-347A-33

Query Match 100.0%; Score 1745; DB 3; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVOFSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGQSSFFS 120

```

; Sequence 1, Application US/08690347
; Patent No. 6383763
; GENERAL INFORMATION:
; APPLICANT: Wallis, Robert S.
; TITLE OF INVENTION: DETECTION OF MYCOBACTERIA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08690,347
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: CASE-02270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; US-08-690-347-1

```

```

Query Match          100.0%; Score 1745; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Db      1  MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60

QY      61  DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
Db      61  DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120

QY      121  DWYSPACGKAGCCTYKWTETLTSELPOWLSANRAVKPTGSAAILGLSMAGSSAMILAAVHP 180
Db      121  DWYSPACGKAGCCTYKWTETLTSELPOWLSANRAVKPTGSAAILGLSMAGSSAMILAAVHP 180

QY      181  QQFIYAGSLALLDPSCGMPSLIGLAMGAGGYKAADMWGPSSDPAWERNNDPTQIIPKL 240
Db      181  QQFIYAGSLALLDPSCGMPSLIGLAMGAGGYKAADMWGPSSDPAWERNNDPTQIIPKL 240

QY      241  VANNTRLWYCGNGTPELGGANI PAEFLNFVRSSNLKFDAYNAAGGHNAVFNPFPNG 300
Db      241  VANNTRLWYCGNGTPELGGANI PAEFLNFVRSSNLKFDAYNAAGGHNAVFNPFPNG 300

QY      301  THSWYWGQALNMGKDLQSSLGAG 325
Db      301  THSWYWGQALNMGKDLQSSLGAG 325

```

RESULT 8

```

US-09-205-426-33
; Sequence 33, Application US/09205426
; Patent No. 6406704
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.

```

```

; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002C4
; CURRENT APPLICATION NUMBER: US/09/205,426
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: 09/095,855
; EARLIER FILING DATE: 1998-06-11
; EARLIER APPLICATION NUMBER: 08/997,362
; EARLIER FILING DATE: 1997-12-23
; EARLIER APPLICATION NUMBER: 08/873,970
; EARLIER FILING DATE: 1997-06-12
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; US-09-205-426-33

```

```

Query Match          100.0%; Score 1745; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Db      1  MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60

QY      61  DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
Db      61  DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120

QY      121  DWYSPACGKAGCCTYKWTETLTSELPOWLSANRAVKPTGSAAILGLSMAGSSAMILAAVHP 180
Db      121  DWYSPACGKAGCCTYKWTETLTSELPOWLSANRAVKPTGSAAILGLSMAGSSAMILAAVHP 180

QY      181  QQFIYAGSLALLDPSCGMPSLIGLAMGAGGYKAADMWGPSSDPAWERNNDPTQIIPKL 240
Db      181  QQFIYAGSLALLDPSCGMPSLIGLAMGAGGYKAADMWGPSSDPAWERNNDPTQIIPKL 240

QY      241  VANNTRLWYCGNGTPELGGANI PAEFLNFVRSSNLKFDAYNAAGGHNAVFNPFPNG 300
Db      241  VANNTRLWYCGNGTPELGGANI PAEFLNFVRSSNLKFDAYNAAGGHNAVFNPFPNG 300

QY      301  THSWYWGQALNMGKDLQSSLGAG 325
Db      301  THSWYWGQALNMGKDLQSSLGAG 325

```

RESULT 9

```

US-09-200-643-33
; Sequence 33, Application US/09200643
; Patent No. 6410720
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002CON
; CURRENT APPLICATION NUMBER: US/09/200,643
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; US-09-200-643-33

```

Query Match 100.0%; Score 1745; DB 4; Length 325;
Best Local Similarity 100.0%; Pred. No. 2.5e-161;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDSRKIRAWGRRLMIGTAAAVVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
Db 1 MTDSRKIRAWGRRLMIGTAAAVVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60

QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFY 120
Db 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFY 120

QY 121 DWSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSWAGSSAMILAAVHP 180
Db 121 DWSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSWAGSSAMILAAVHP 180

QY 181 QOFTYAGLSALLDPSQGMGSLIGLAWGDAGGYKAADMWGPSSDPPOWERNDPTQQIPKL 240
Db 181 QOFTYAGLSALLDPSQGMGSLIGLAWGDAGGYKAADMWGPSSDPPOWERNDPTQQIPKL 240

QY 241 VANNTRLMVYCGNGTNPNEGANIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFPPNG 300
Db 241 VANNTRLMVYCGNGTNPNEGANIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFPPNG 300

QY 301 THSWEYGAQINAMKGDLOSSLGAG 325
Db 301 THSWEYGAQINAMKGDLOSSLGAG 325

RESULT 10
US-08-107-676-29
; Sequence 29, Application US/08107676
; Patent No. 5955356
; GENERAL INFORMATION:
; APPLICANT: Content, Jean
; APPLICANT: DeBruyn, Jacqueline
; TITLE OF INVENTION: Mycobacterium Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding Them for Diagnosis and Control of
; TITLE OF INVENTION: Tuberculosis
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESS: Merchant & Gould
; STREET: 3100 No. 5955356west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/107,676
; FILING DATE: 04-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP92/00268
; FILING DATE: 07-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91400388.4
; FILING DATE: 14-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 8076.89USWO
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
ORIGINAL SOURCE: Mycobacterium bovis
ORGANISM: Mycobacterium bovis
IMMEDIATE SOURCE:
CLONE: Antigen 85B protein sequence from
CLONE: alpha-antigen of M.bovis
US-08-107-676-29

Query Match 99.0%; Score 1728; DB 2; Length 325;
Best Local Similarity 99.1%; Pred. No. 1.1e-159;
Matches 322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MTDSRKIRAWGRRLMIGTAAAVVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
Db 1 MTDSRKIRAWGRRLMIGTAAAVVLPGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60

QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFY 120
Db 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFY 120

QY 121 DWSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSWAGSSAMILAAVHP 180
Db 121 DWSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSWAGSSAMILAAVHP 180

QY 181 QOFTYAGLSALLDPSQGMGSLIGLAWGDAGGYKAADMWGPSSDPPOWERNDPTQQIPKL 240
Db 181 QOFTYAGLSALLDPSQGMGSLIGLAWGDAGGYKAADMWGPSSDPPOWERNDPTQQIPKL 240

QY 241 VANNTRLMVYCGNGTNPNEGANIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFPPNG 300
Db 241 VANNTRLMVYCGNGTNPNEGANIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFPPNG 300

QY 301 THSWEYGAQINAMKGDLOSSLGAG 325
Db 301 THSWEYGAQINAMKGDLOSSLGAG 325

RESULT 11
US-08-997-080-35
; Sequence 35, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESS: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565

TELEFAX: 206-269-0563

TELEX:

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:

LENGTH: 323 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-997-080-35

Query Match 97.8%; Score 1706; DB 2; Length 323;
 Best Local Similarity 98.5%; Pred. No. 1.5e-157;
 Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

```

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
DB 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
QY 121 DWSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
DB 121 DWSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
QY 181 QQFIYAGLSALLDPSQGMPSLIGLWAGDAGGYKAADMWGPSSDDPAWERNPTQQIPKL 240
DB 181 QQFIYAGLSALLDPSQGMG--LIGLWAGDAGGYKAADMWGPSSDDPAWERNPTQQIPKL 238
QY 241 VANNTRLVYCGNGTNPNEGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVNFPPNG 300
DB 239 VANNTRLVYCGNGTNPNEGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVNFPPNG 298
QY 301 THSWEYWGALNAMKGDLOSSLGAG 325
DB 299 THSWEYWGALNAMKGDLOSSLGAG 323

```

RESULT 12

US-08-997-362-35

Sequence 35, Application US/08997362

Patent No. 5985287

GENERAL INFORMATION:

APPLICANT: Tan, Paul

APPLICANT: Hiyama, Jun

APPLICANT: Visser, Elizabeth

APPLICANT: Skinner, Margot

APPLICANT: Scott, Linda

APPLICANT: Prestidge, Ross

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR

TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS

NUMBER OF SEQUENCES: 194

CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/997,362

FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970

FILING DATE: June 12, 1997

APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347

FILING DATE: August 29, 1996

ATTORNEY/AGENT INFORMATION:

NAME: Sleath, Janet

REGISTRATION NUMBER: 37,007

REFERENCE/DOCKET NUMBER: 11000.1002c2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 206-269-0565

TELEFAX: 206-269-0563

TELEX:

INFORMATION FOR SEQ ID NO: 35:

SEQUENCE CHARACTERISTICS:

LENGTH: 323 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-997-362-35

Query Match

Best Local Similarity 97.8%; Score 1706; DB 2; Length 323;

Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

```

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
DB 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
QY 121 DWSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
DB 121 DWSPACGKAGCQTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMILAAAYHP 180
QY 181 QQFIYAGLSALLDPSQGMPSLIGLWAGDAGGYKAADMWGPSSDDPAWERNPTQQIPKL 240
DB 181 QQFIYAGLSALLDPSQGMG--LIGLWAGDAGGYKAADMWGPSSDDPAWERNPTQQIPKL 238
QY 241 VANNTRLVYCGNGTNPNEGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVNFPPNG 300
DB 239 VANNTRLVYCGNGTNPNEGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVNFPPNG 298
QY 301 THSWEYWGALNAMKGDLOSSLGAG 325
DB 299 THSWEYWGALNAMKGDLOSSLGAG 323

```

RESULT 13

US-08-873-970-35

Sequence 35, Application US/08873970

Patent No. 6001361

GENERAL INFORMATION:

APPLICANT: Tan, Paul

APPLICANT: Hiyama, Jun

APPLICANT: Visser, Elizabeth

APPLICANT: Skinner, Margot

APPLICANT: Scott, Linda

APPLICANT: Prestidge, Ross

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR

TREATMENT AND DIAGNOSIS OF MYCOBACTERIAL INFECTIONS

NUMBER OF SEQUENCES: 106

CORRESPONDENCE ADDRESS:

ADDRESSEE: Law Offices of Ann W. Speckman

STREET: 2601 Elliott Avenue, Suite 4185

CITY: Seattle

STATE: WA

COUNTRY: USA

ZIP: 98121

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/095,855
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA: 08/705,347
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
APPLICATION NUMBER: 08/873,970
FILING DATE: 12-JUN-1997
APPLICATION NUMBER: 08/997,362
FILING DATE: 23-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 323 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-C95-855-35

Query Match 97.8%; Score 1706; DB 3; Length 323;
Best Local Similarity 98.5%; Pred. No. 1.5e-157;
Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
DB 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
QY 121 DWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
DB 121 DWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
QY 181 QCFIYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGFSSDPAWERNPTQIQPKL 240
DB 181 QCFIYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGFSSDPAWERNPTQIQPKL 238
QY 241 VANNTLWYCGNGTNEELGGANIPAEFLENFVRSSNLKFDAYNAAGHNAVFPNPPNG 300
DB 239 VANNTLWYCGNGTNEELGGANIPAEFLENFVRSSNLKFDAYNAAGHNAVFPNPPNG 298
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
DB 299 THSWEYGAQLNAMKGDLOSSLGAG 323

Query Match 97.8%; Score 1706; DB 3; Length 323;
Best Local Similarity 98.5%; Pred. No. 1.5e-157;
Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
DB 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
QY 121 DWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
DB 121 DWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
QY 181 QCFIYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGFSSDPAWERNPTQIQPKL 240
DB 181 QCFIYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGFSSDPAWERNPTQIQPKL 238
QY 241 VANNTLWYCGNGTNEELGGANIPAEFLENFVRSSNLKFDAYNAAGHNAVFPNPPNG 300
DB 239 VANNTLWYCGNGTNEELGGANIPAEFLENFVRSSNLKFDAYNAAGHNAVFPNPPNG 298
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
DB 299 THSWEYGAQLNAMKGDLOSSLGAG 323

RESULT 15
US-08-705-347A-35
; Sequence 35, Application US/08705347A
; Patent No. 6284255
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margo
; APPLICANT: Scott, Linda
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT AND

SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/873,970
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: 08/705,347
APPLICATION NUMBER: 08/705,347
FILING DATE: 29-AUG-1996
ATTORNEY/AGENT INFORMATION:
NAME: Sleath, Janet
REGISTRATION NUMBER: 37,007
REFERENCE/DOCKET NUMBER: 11000.1002C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-269-0565
TELEFAX: 206-269-0563
TELEX:
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 323 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-873-970-35

Query Match 97.8%; Score 1706; DB 3; Length 323;
Best Local Similarity 98.5%; Pred. No. 1.5e-157;
Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
DB 61 DIKVFQSGGNNSPAVYLLDGLRAODDYNWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
QY 121 DWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
DB 121 DWYSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAAGLSMAGSSAMILAAVHP 180
QY 181 QCFIYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGFSSDPAWERNPTQIQPKL 240
DB 181 QCFIYAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGFSSDPAWERNPTQIQPKL 238
QY 241 VANNTLWYCGNGTNEELGGANIPAEFLENFVRSSNLKFDAYNAAGHNAVFPNPPNG 300
DB 239 VANNTLWYCGNGTNEELGGANIPAEFLENFVRSSNLKFDAYNAAGHNAVFPNPPNG 298
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
DB 299 THSWEYGAQLNAMKGDLOSSLGAG 323

RESULT 14
US-09-095-855-35
; Sequence 35, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margo
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121

; TITLE OF INVENTION: DIAGNOSIS OF MYCOBACTERIAL INFECTIONS
 ; NUMBER OF SEQUENCES: 55
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Speckman Picard PLLC
 ; STREET: 2601 Elliott Avenue, Suite 4185
 ; CITY: Seattle
 ; STATE: WA
 ; COUNTRY: USA
 ; ZIP: 98121
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/705,347A
 ; FILING DATE: 28-AUG-1996
 ; CLASSIFICATION: 424
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Sleath, Janet
 ; REGISTRATION NUMBER: 37,007
 ; REFERENCE/DOCKET NUMBER: 11000.1002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 206.269.0565
 ; TELEFAX: 206.269.0563
 ; INFORMATION FOR SEQ ID NO: 35:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 323 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-705-347A-35

Query Match 97.8%; Score 1706; DB 3; Length 323;
 Best Local Similarity 98.5%; Pred. No. 1.5e-157;
 Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
 QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
 Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
 QY 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSFYS 120
 Db 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSFYS 120
 QY 121 DWYSPACGKAGCQTYKWETLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
 Db 121 DWYSPACGKAGCQTYKWETLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
 QY 181 QQFIYAGSLSALLDPSQGMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNPTQOIPKL 240
 Db 181 QQFIYAGSLSALLDPSQGMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNPTQOIPKL 238
 QY 241 VANNTRLWYVCGTGNELGGANIPAEFLNFVRSSNLKFDQAYNAAGHNAVFNPFG 300
 Db 239 VANNTRLWYVCGTGNELGGANIPAEFLNFVRSSNLKFDQAYNAAGHNAVFNPFG 298
 QY 301 THSWEYWGQALNMGDLOSSLGAG 325
 Db 299 THSWEYWGQALNMGDLOSSLGAG 323

RESULT 16
 US-09-324-542-35
 ; Sequence 35, Application US/09324542
 ; Patent No. 6328978
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Tan, Paul L.J.
 ; APPLICANT: Prestidge, Ross
 ; TITLE OF INVENTION: Methods and Compounds for the Treatment
 ; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders

; FILE REFERENCE: 11000.1007c1
 ; CURRENT APPLICATION NUMBER: US/09/324,542
 ; CURRENT FILING DATE: 1999-06-02
 ; EARLIER APPLICATION NUMBER: US 08/997,080
 ; EARLIER FILING DATE: 1997-12-23
 ; NUMBER OF SEQ ID NOS: 194
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 35
 ; LENGTH: 323
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium bovis
 ; US-09-324-542-35

Query Match 97.8%; Score 1706; DB 4; Length 323;
 Best Local Similarity 98.5%; Pred. No. 1.5e-157;
 Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;
 QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
 Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPGVLGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
 QY 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSFYS 120
 Db 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSFYS 120
 QY 121 DWYSPACGKAGCQTYKWETLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
 Db 121 DWYSPACGKAGCQTYKWETLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
 QY 181 QQFIYAGSLSALLDPSQGMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNPTQOIPKL 240
 Db 181 QQFIYAGSLSALLDPSQGMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNPTQOIPKL 238
 QY 241 VANNTRLWYVCGTGNELGGANIPAEFLNFVRSSNLKFDQAYNAAGHNAVFNPFG 300
 Db 239 VANNTRLWYVCGTGNELGGANIPAEFLNFVRSSNLKFDQAYNAAGHNAVFNPFG 298
 QY 301 THSWEYWGQALNMGDLOSSLGAG 325
 Db 299 THSWEYWGQALNMGDLOSSLGAG 323

RESULT 17
 US-09-205-426-35
 ; Sequence 35, Application US/09205426
 ; Patent No. 6406704
 ; GENERAL INFORMATION:
 ; APPLICANT: Watson, James D.
 ; APPLICANT: Tan, Paul L. J.
 ; TITLE OF INVENTION: Compounds and Methods for Treatment and
 ; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
 ; FILE REFERENCE: 11000.1002c4
 ; CURRENT APPLICATION NUMBER: US/09/205,426
 ; CURRENT FILING DATE: 1998-12-04
 ; EARLIER APPLICATION NUMBER: 09/095,855
 ; EARLIER FILING DATE: 1998-06-11
 ; EARLIER APPLICATION NUMBER: 08/997,362
 ; EARLIER FILING DATE: 1997-12-23
 ; EARLIER APPLICATION NUMBER: 08/873,970
 ; EARLIER FILING DATE: 1997-06-12
 ; EARLIER APPLICATION NUMBER: 08/705,347
 ; EARLIER FILING DATE: 1996-08-29
 ; NUMBER OF SEQ ID NOS: 208
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 35
 ; LENGTH: 323
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium bovis
 ; US-09-205-426-35

Query Match 97.8%; Score 1706; DB 4; Length 323;
 Best Local Similarity 98.5%; Pred. No. 1.5e-157;
 Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

```

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSFSGMR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSFSGMR 60
QY 61 DIKQVQFQSGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
Db 61 DIKQVQFQSGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
QY 121 DWTSPACGKAGCCTYKWEITLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAAYHP 180
Db 121 DWTSPACGKAGCCTYKWEITLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAAYHP 180
QY 181 QQFIYAGSLALLDPSQCMGSLIGLAWDAGGYKAADMWGPSSDPAWERNPTQOI PKL 240
Db 181 QQFIYAGSLALLDPSQCMG--LIGLAWDAGGYKAADMWGPSSDPAWERNPTQOI PKL 238
QY 241 VANNTRLVVYCGNGTNPNEGANIPAEFFLENFVRSSNLKFQDAYNAAGGHNAVENFPFNG 300
Db 239 VANNTRLVVYCGNGTNPNEGANIPAEFFLENFVRSSNLKFQDAYKPAGGHNAVENFPFNG 298
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
Db 299 THSWEYGAQLNAMKGDLOSSLGAG 323

RESULT 18
US-09-200-643-35
; Sequence 35, Application US/09200643
; Patent No. 6410720
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Compounds and Methods for Treatment and
; TITLE OF INVENTION: Diagnosis of Mycobacterial Infections
; FILE REFERENCE: 11000.1002CON
; CURRENT APPLICATION NUMBER: US/09/200,643
; CURRENT FILING DATE: 1998-11-05
; EARLIER APPLICATION NUMBER: 08/705,347
; EARLIER FILING DATE: 1996-08-29
; NUMBER OF SEQ ID NOS: 55
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-09-200-643-35

Query Match 97.8%; Score 1706; DB 4; Length 323;
Best Local Similarity 98.5%; Pred. No. 1.5e-157; Indels 2; Gaps 1;
Matches 320; Conservative 0; Mismatches 3;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSFSGMR 60
Db 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSFSGMR 60
QY 61 DIKQVQFQSGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
Db 61 DIKQVQFQSGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYS 120
QY 121 DWTSPACGKAGCCTYKWEITLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAAYHP 180
Db 121 DWTSPACGKAGCCTYKWEITLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAAYHP 180
QY 181 QQFIYAGSLALLDPSQCMGSLIGLAWDAGGYKAADMWGPSSDPAWERNPTQOI PKL 240
Db 181 QQFIYAGSLALLDPSQCMG--LIGLAWDAGGYKAADMWGPSSDPAWERNPTQOI PKL 238
QY 241 VANNTRLVVYCGNGTNPNEGANIPAEFFLENFVRSSNLKFQDAYNAAGGHNAVENFPFNG 300
Db 239 VANNTRLVVYCGNGTNPNEGANIPAEFFLENFVRSSNLKFQDAYKPAGGHNAVENFPFNG 298
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325

Query Match 93.0%; Score 1623; DB 3; Length 307;
Best Local Similarity 98.4%; Pred. No. 1.6e-149; Indels 2; Gaps 1;
Matches 304; Conservative 0; Mismatches 3;

QY 17 IGTAATAVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSFSGMRDIKVOFQSGNNSPAV 76
Db 1 IGTAATAVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSFSGMRDIKVOFQSGNNSPAV 60
QY 77 YLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYSWYSPACGKAGCCTYK 136
Db 61 YLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFYSWYSPACGKAGCCTYK 120
QY 137 WETFLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAAYHPQQFIYAGSLALLDPS 196
Db 121 WETFLTSELPOWLSANRAVKPTGSAALGLSMAGSSAMILAAAYHPQQFIYAGSLALLDPS 180
QY 197 QMGPSLIGLAWDAGGYKAADMWGPSSDPAWERNPTQOI PKLVANNTRLVVYCGNGT 256
Db 197 QMGPSLIGLAWDAGGYKAADMWGPSSDPAWERNPTQOI PKLVANNTRLVVYCGNGT 256

```


Db 181 QCMG--LIGLAMDAGGYKAADWGFSSDPWRNDPTQIIPKLVANNTRLVYCGNGTP 238
QY 257 NELGANTPABPLENFVRSSNLKFDQAYNAAGHNAVFNPNGTHSWEYGAQINAMKG 316
Db 239 NELGANTPABPLENFVRSSNLKFDQAYKAPAGHNAVFNPNGTHSWEYGAQINAMKG 298
QY 317 DQSSLGAG 325
Db 299 DQSSLGAG 307

RESULT 20
US-09-051-755-16
; Sequence 16, Application US/09051755
; Patent No. 6506389
; GENERAL INFORMATION:
; APPLICANT: LEER, Robert J
; APPLICANT: POWWELS, Pieter H
; APPLICANT: CONWAY, Patricia L
; TITLE OF INVENTION: NOVEL ADHERENCE FACTORS OF NON PATHOGENIC
; TITLE OF INVENTION: MICROORGANISMS AND APPLICATIONS THEREOF FOR SCREENING
; TITLE OF INVENTION: MICROORGANISMS FOR SPECIFIC PROBIOTIC PROPERTIES; NOVEL
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS AND FOOD ADDITIVES C
; FILE REFERENCE: LEER et al. 09/051.755
; CURRENT APPLICATION NUMBER: US/09/051.755
; EARLIER FILING DATE: 1998-09-30
; EARLIER APPLICATION NUMBER: PCT/NL95/00367
; EARLIER FILING DATE: 1995-10-20
; EARLIER APPLICATION NUMBER: PCT/NL96/00409
; EARLIER FILING DATE: 1996-10-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 16
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: aligned aa
; OTHER INFORMATION: sequence of Mycobacterium Mtu85B
; NAME/KEY: UNSURE
; LOCATION: (1)..(364)
; OTHER INFORMATION: Xaa is any amino acid.
US-09-051-755-16

Query Match 92.4%; Score 1612; DB 4; Length 364;
Best Local Similarity 92.6%; Pred. No. 2.4e-148;
Matches 325; Conservative 0; Mismatches 0; Indels 26; Gaps 12;
QY 1 MTDVSRKIRA---WGRRLMIGTAAAVLPLGLVGLAGGATAGAFSRPG-LPVEYLQVPS 55
Db 3 MTDVSRKIRAXXXWGRRLMIGTAAAVLPLGLVGLAGGATAGAFSRPG-LPVEYLQVPS 62
QY 56 P-SMGDRDKV-QFQSGGNNSPAVYLLDGLRAQDDY--NGWDINTPAFEWYQSG-LSIYM 110
Db 63 PSMGDRDKVQFQSGGNNSPAVYLLDGLRAQDDYXXNGWDINTPAFEWYQSGXLSIYM 122
QY 111 PVGGQSSFYSDWYSPACGAGCTYKWT-FTLSELPOWLSANRAVKPTGSAAGI----- 165
Db 123 PVGGQSSFYSDWYSPACGAGCTYKWTFTLSELPOWLSANRAVKPTGSAAGI----- 182
QY 166 --SMAGSS---AMLAAYHPQFI---YAGSLSALDPSQGMGPSLIGLAMDAGGYKAA 217
Db 183 XSMAGSSXXXAMILAAAYHPQFIXXXAGSLSALDPSQGMGPSLIGLAMDAGGYKAA 242
QY 218 DMWGSSDPWRNDPTQIIPKLVANNTRLVYCGNGTNEELGGANI PAEFLFNVRSSN 277
Db 243 DMWGSSDPWRNDPTQIIPKLVANNTRLVYCGNGTNEELGGANI PAEFLFNVRSSN 302
QY 278 LKFDAYNAAGHNAVFNPNGTHSWE--YGAQINAMKGLQSSSL--GAG 325
Db 303 LKFDAYNAAGHNAVFNPNGTHSWEYXXYGAQINAMKGLDQSSSLXGAG 353

RESULT 21
US-08-107-676-30
; Sequence 30, Application US/08107676
; Patent No. 5955356
; GENERAL INFORMATION:
; APPLICANT: Content, Jean
; APPLICANT: DeWit, Lucas
; APPLICANT: DeBruyn, Jacqueline
; TITLE OF INVENTION: Mycobacterium Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding Them for Diagnosis and Control of
; TITLE OF INVENTION: Tuberculosis
; NUMBER OF SEQUENCES: 38
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merchant & Gould
; STREET: 3100 No. 5955356west Center
; CITY: Minneapolis
; STATE: MN
; COUNTRY: USA
; ZIP: 55402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; APPLICATION NUMBER: US/08/107,676
; FILING DATE: 04-OCT-1993
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP92/00268
; FILING DATE: 07-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 91400388.4
; FILING DATE: 14-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Kowalchuk, Katherine M.
; REGISTRATION NUMBER: 36,848
; REFERENCE/DOCKET NUMBER: 8076.89USWO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 612-332-5300
; TELEFAX: 612-332-9081
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 325 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Mycobacterium kansasii
; IMMEDIATE SOURCE:
; CLONE: Partial protein sequence from antigen 85B
; CLONE: from M.kansasii
US-08-107-676-30

Query Match 92.2%; Score 1609; DB 2; Length 325;
Best Local Similarity 89.8%; Pred. No. 4e-148;
Matches 291; Conservative 20; Mismatches 13; Indels 0; Gaps 0;
QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPLGLVGLAGGATAGAFSRPG-LPVEYLQVPSMGR 60
Db 1 MTDVSGKIRAWGRRLLVGAARAAALPLGLVGLAGGATAGAFSRPG-LPVEYLQVPSAAMGR 60
QY 61 DIKVOFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIYMPVGGQSSFY 120
Db 61 SIKVOFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIYMPVGGQSSFY 120
QY 121 DWYSPACGAGCTYKWTFTLSELPOWLSANRAVKPTGSAAGI-SMAGSSAMILAAAYHP 180
Db 121 DWYSPACGAGCTYKWTFTLSELPOWLSANRAVKPTGSAAGI-SMAGSSAMILSVYHP 180
QY 181 QOFTYAGSLSAILDPSQGMGPSLIGLAMDAGGYKAADWGFSSDPWRNDPTQIIPKL 240

```

Db 181 QQFIYAGSLALMDPSQGMPSLIGLWGDAGGYKASDMWGPSSDPWQNRDPSLHPEL 240
QY 241 VANNTRLWVYCGTGNELGANIPAEFFLENFVRSSNLKFDQDAYNAAAGGNVNFPPNG 300
Db 241 VANNTRLWVYCGTGNELGANIPAEFFLENFVRSSNLKFDQDAYNAAAGGNVNFPPNG 300
QY 301 THSWEYWGAGLQINAMKGLQSSLSGA 324
Db 301 THSWEYWGAGLQINAMKGLQSSLSGA 324

RESULT 22
US-08-508-761B-32
; Sequence 32, Application US/08508761B
; Patent No. 6027920
; GENERAL INFORMATION:
; APPLICANT: Joliff, Gwennael
; APPLICANT: Guyonvarch, Arnel
; APPLICANT: Purification, Relano
; APPLICANT: Duchiron, Francis
; APPLICANT: Renaud, Michel
; TITLE OF INVENTION: System for Protein Expression and
; SECRETION Especially in Corynebacteria
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Jacobson, Price, Holman & Stern, PLLC
; STREET: 400 Seventh St. N.W.
; CITY: Washington D.C.
; COUNTRY: U.S.A.
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/508,761B
; FILING DATE: 31-JUL-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 91/09652
; FILING DATE: 29-JUL-1991
; APPLICATION NUMBER: FR 91/09870
; FILING DATE: 02-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Player, William E.
; REGISTRATION NUMBER: 31,409
; REFERENCE/DOCKET NUMBER: P58525NA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638-6666
; TELEFAX: (202) 393-5350
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: M. kansaii
US-08-508-761B-32

Query Match 88.0%; Score 1536; DB 3; Length 309;
Best Local Similarity 89.3%; Pred. No. 4.6e-141;
Matches 277; Conservative 19; Mismatches 12; Indels 0; Gaps 0;

QY 17 ICTAAAVLPGLVGLAGGATAGAFSRPGLPVEYLQVPSMGDR:KVQFQSGGNNSPAV 76
Db 1 VGGAAAAALPGLVGLAGGATAGAFSRPGLPVEYLQVPSMGDR:KVQFQSGGNNSPAV 60
QY 77 YLLDGLRAQDYGNDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCOTYK 136
Db 61 YLLDGLRAQDYGNDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCOTYK 120

```

```

QY 137 WETFLTSELQWL SANRAVKPTGSAALGLSMAGSSAMILAAHYHPQOPIYAGSLALDPS 196
Db 121 WETFLTSELQWL SANRAVKPTGSAALGLSMAGSSAMILAAHYHPQOPIYAGSLALDPS 180
QY 197 QGMGPSLIGLWGDAGGYKASDMWGPSSDPWQNRDPSLHPEL VANNTRLWVYCGTGN 256
Db 181 QGMGPSLIGLWGDAGGYKASDMWGPSSDPWQNRDPSLHPEL VANNTRLWVYCGTGN 240
QY 257 NEELGGANIPAEFFLENFVRSSNLKFDQDAYNAAAGGNVNFPPNGTHSWEYWGAGLQINAMK 316
Db 241 SELGGANVPAEFFLENFVRSSNLKFDQDAYNAAAGGNVNFPPNGTHSWEYWGAGLQINAMK 300
QY 317 DLOSSLGA 324
Db 301 DLOSSLGA 308

RESULT 23
US-08-690-347-2
; Sequence 2, Application US/08690347
; Patent No. 6383763
; GENERAL INFORMATION:
; APPLICANT: Wallis, Robert S.
; TITLE OF INVENTION: DETECTION OF MYCOBACTERIA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,347
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: CASE-02270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
US-08-690-347-2

Query Match 87.2%; Score 1522; DB 4; Length 330;
Best Local Similarity 84.9%; Pred. No. 1.1e-139;
Matches 275; Conservative 24; Mismatches 25; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGATAGAFSRPGLPVEYLQVPSMGDR 60
Db 1 MTDLSEKVRANRGLLVLGAAAVLPGLVGLAGGATAGAFSRPGLPVEYLQVPSMGDR 60
QY 61 DIKVPQSGGNNSPAVYLLDGLRAQDYGNDINTPAFEWYQSGLSIVMPVGGQSSFY 120
Db 61 DIKVPQSGGNNSPAVYLLDGLRAQDYGNDINTPAFEWYQSGLSIVMPVGGQSSFY 120
QY 121 DWYSPACGKAGCOTYKWTFLTSELQWL SANRAVKPTGSAALGLSMAGSSAMILAAHYHP 180
Db 121 DWYSPACGKAGCOTYKWTFLTSELQWL SANRAVKPTGSAALGLSMAGSSAMILAAHYHP 180

```

QY 181 QQFIVAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPANERNDPTQOIPKL 240
Db 181 DOFIVAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPANERNDPTQOIPKL 240
QY 241 VANNTRLVVYCGNGTNPNEGANIPAEFLENFVRSSNLKFODAYNAAGHNAVFNPNG 300
Db 241 VGHNTRLVVYCGNGTNPSELGANNPAEFLENFVRSSNLKFODAYNGAGHNAVFNPNG 300
QY 301 THSWEYGAQLNANKGDIQSSIGA 324
Db 301 THSWEYGAQLNANKPDLOQTIGA 324

RESULT 24

US-08-311-731A-83
; Sequence 83, Application US/08311731A
; Patent No. 6583266
; GENERAL INFORMATION:
; APPLICANT: SMITH, DOUGLAS
; APPLICANT: MAO, JEN-I
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 411
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WOLF GREENFIELD & SACKS, P.C.
; STREET: 600 ATLANTIC AVENUE
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,731A
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: GATES, EDWARD R.
; REGISTRATION NUMBER: 31,616
; REFERENCE/DOCKET NUMBER: C0044/7125
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/720-3500
; TELEFAX: 617/720-2441
; INFORMATION FOR SEQ ID NO: 83:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: MYCOBACTERIUM LEPRAE
US-08-311-731A-83

Query Match 83.4%; Score 1455; DB 4; Length 327;
Best Local Similarity 83.0%; Pred. No. 3 6e-133;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;
QY 1 MIDVSKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSPMGR 60
Db 1 MIDVSKIRAWGRLLVVG-AAATLPLSLISLAGGAATASAFSRPGLPVEYLQVPSEAMGR 58
QY 61 DIKVOFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
Db 59 TIKVOFQNGGNSPAVYLLDGLRAQDDYNGWDINTSAFEWYQSGLSVMPVGGSSFFYS 118
QY 121 DWYSPACGKAGCTYKWTFTLTSELPOWLSANRAVKPTGSAAGISLMSAGSSAMIIAAVHP 180

Db 119 DWYSPACGKAGCTYKWTFTLTSELPKWLSANRSVKSTGSAVVGSLSMAGSSALIIAAVHP 178
QY 181 QQFIVAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPANERNDPTQOIPKL 240
Db 179 DOFIVAGSLALLDPSQGMPSLIGLAMGDAGGYKAADMWGPSSDPANERNDPTQOIPKL 238
QY 241 VANNTRLVVYCGNGTNPNEGANIPAEFLENFVRSSNLKFODAYNAAGHNAVFNPNG 300
Db 239 VANNTRLVVYCGNGTNPSELGANNPAEFLENFVRSSNLKFODAYNGAGHNAVFNPNG 298
QY 301 THSWEYGAQLNANKGDIQSSIGA 324
Db 299 THSWEYGAQLNANKPDLOQTIGA 322

RESULT 25

US-08-997-080-31
; Sequence 31, Application US/08997080
; Patent No. 5968524
; GENERAL INFORMATION:
; APPLICANT: WATSON, JAMES D.
; APPLICANT: TAN, PAUL L.J.
; TITLE OF INVENTION: METHODS AND COMPOUNDS FOR THE TREATMENT OF IMMUNOLOGICALLY-
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,080
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1007
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-997-080-31

Query Match 83.3%; Score 1454; DB 2; Length 327;
Best Local Similarity 83.0%; Pred. No. 4 5e-133;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;
QY 1 MIDVSKIRAWGRRLMIGTAAAVLPGVLGAGGAATAGAFSRPGLPVEYLQVPSPMGR 60
Db 1 MIDVSKIRAWGRLLVVG-AAATLPLSLISLAGGAATASAFSRPGLPVEYLQVPSEAMGR 58
QY 61 DIKVOFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFFYS 120
Db 59 TIKVOFQNGGNSPAVYLLDGLRAQDDYNGWDINTSAFEWYQSGLSVMPVGGSSFFYS 118
QY 121 DWYSPACGKAGCTYKWTFTLTSELPOWLSANRAVKPTGSAAGISLMSAGSSAMIIAAVHP 180

Db 119 DWTSPACGKAGCTTYKWEITELTSELPAKWSANRSVKSTGSAVVGSLMAGSSALILAAHP 178
QY 181 QQFIYAGSLALDPSQMGPSLIGLAMGAGGYKAADMWGPSSDPAWERNDPQQIPKL 240
Db 179 DQFIYAGSLALMDSSQIEPQLIGLAMGAGGYKAADMWGPSPNDPAWQRNDPILQAGKL 238
QY 241 VANNTRLWVYCGNGTPELGGANIPAEFLENFVRSNLFKQDAYNAAGGHNAVENFPFNG 300
Db 239 VANNTHLWVYCGNGTPELGGTNVPAEFLENFVHGSNLKFDQAYNGAGGHNAVENLNADG 298
QY 301 THSWEYGAQNLNMGDLQSSLGA 324
Db 299 THSWEYGAQNLNMGKPDQLNTLMA 322

RESULT 26
US-08-997-362-31
; Sequence 31, Application US/08997362
; Patent No. 5985287
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 194
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/997,362
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/873,970
; FILING DATE: June 12, 1997
; APPLICATION NUMBER: U.S. Patent Application No. 5985287 08/705,347
; FILING DATE: August 29, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 11000.1002c2
; REFERENCE/DOCKET NUMBER: 11000.1002c2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-997-362-31

Query Match 83.3%; Score 1454; DB 2; Length 327;
Best Local Similarity 83.0%; Pred. No. 4.5e-133;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;
QY 1 MTDVSKIRAWGRRLMIGTAAAVLPLGLVLAGGAATAGAFSRPGLFVEYLQVPSPMGR 60

Db 1 MTDVSKIRAWGRLLVG--AAATLPSLISLAGGAATASAFSRPGLFVEYLQVPSEAMGR 58
QY 61 DIKVFQSGGNNPARYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVNPFVQSGSFYS 120
Db 59 TIKVFQFQNGGSPAVYLLDGLRAQDDYNGWDINTSAFEWYQSGLSIVNPFVQSGSFYS 118
QY 121 DWTSPACGKAGCTTYKWEITELTSELPAKWSANRSVKSTGSAVVGSLMAGSSALILAAHP 180
Db 119 DWTSPACGKAGCTTYKWEITELTSELPAKWSANRSVKSTGSAVVGSLMAGSSALILAAHP 178
QY 181 QQFIYAGSLALDPSQMGPSLIGLAMGAGGYKAADMWGPSSDPAWERNDPQQIPKL 240
Db 179 DQFIYAGSLALMDSSQIEPQLIGLAMGAGGYKAADMWGPSPNDPAWQRNDPILQAGKL 238
QY 241 VANNTRLWVYCGNGTPELGGANIPAEFLENFVRSNLFKQDAYNAAGGHNAVENFPFNG 300
Db 239 VANNTHLWVYCGNGTPELGGTNVPAEFLENFVHGSNLKFDQAYNGAGGHNAVENLNADG 298
QY 301 THSWEYGAQNLNMGDLQSSLGA 324
Db 299 THSWEYGAQNLNMGKPDQLNTLMA 322

RESULT 27
US-08-873-970-31
; Sequence 31, Application US/08873970
; Patent No. 6001361
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Hiyama, Jun
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Scott, Linda
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR
; MYCOBACTERIAL INFECTIONS
; NUMBER OF SEQUENCES: 106
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/873,970
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-873-970-31

```
Query Match      83.3%; Score 1454; DB 3; Length 327;
Best Local Similarity 83.0%; Pred. No. 4.5e-133;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;

QY 1 MTDVSKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
Db 1 MIDVSKIRAWGRWLLVG--AAATLPSLISLAGAATASAFSPGLPVEYLQVPSMGR 58

QY 61 DIKVFQSGGNNSPAYLLDGLRAQDDYNGWDINTPAFEWYQSLIVMPVGGQSSFFS 120
Db 59 TIKVFQNGGNSPAVYLLDGLRAQDDYNGWDINTSAFEWYQSLIVMPVGGQSSFFS 118

QY 121 DWYSPACGAGCCTYKWTFTLSELQWLSANRAVPTGSAALGLSMAGSSAMILAAHP 180
Db 119 DWYSPACGAGCCTYKWTFTLSELQWLSANRAVPTGSAALGLSMAGSSAMILAAHP 178

QY 181 QOFIYAGSLISALIDPSQGMPSLIGLAWGDAGGYKAADMWGPSSDPAWERNDPTQIPL 240
Db 179 DQFIYAGSLISALMDSSQGIPEQLIGLAWGDAGGYKAADMWGPSPNDPAWRNDPILQAGKL 238

QY 241 VANNTLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDQDAYNAGGHNAVFNFPNG 300
Db 239 VANNTLWVYCGNGTPELGGTNPVPAEFLNFVHGSNLKFDQDAYNAGGHNAVFNLNADG 298

QY 301 THSWEYWGALNMAKGDLOSSLGA 324
Db 299 THSWEYWGALNMAKDPDLQNTLMA 322

RESULT 29
US-09-095-855-31
; Sequence 31, Application US/09095855
; Patent No. 6160093
; GENERAL INFORMATION:
; APPLICANT: Tan, Paul
; APPLICANT: Visser, Elizabeth
; APPLICANT: Skinner, Margot
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Compounds and Methods for
; TITLE OF INVENTION: Treatment and Diagnosis of Mycobacterial Infections
; NUMBER OF SEQUENCES: 208
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,855
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/705,347
; FILING DATE: 29-AUG-1996
; APPLICATION NUMBER: 08/873,970
; FILING DATE: 12-JUN-1997
; APPLICATION NUMBER: 08/997,362
; FILING DATE: 23-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1002c3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:

; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-095-855-31

Query Match      83.3%; Score 1454; DB 3; Length 327;
Best Local Similarity 83.0%; Pred. No. 4.5e-133;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;

QY 1 MTDVSKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSPGLPVEYLQVPSMGR 60
Db 1 MIDVSKIRAWGRWLLVG--AAATLPSLISLAGAATASAFSPGLPVEYLQVPSMGR 58

QY 61 DIKVFQSGGNNSPAYLLDGLRAQDDYNGWDINTPAFEWYQSLIVMPVGGQSSFFS 120
Db 59 TIKVFQNGGNSPAVYLLDGLRAQDDYNGWDINTSAFEWYQSLIVMPVGGQSSFFS 118

QY 121 DWYSPACGAGCCTYKWTFTLSELQWLSANRAVPTGSAALGLSMAGSSAMILAAHP 180
Db 119 DWYSPACGAGCCTYKWTFTLSELQWLSANRAVPTGSAALGLSMAGSSAMILAAHP 178

QY 181 QOFIYAGSLISALIDPSQGMPSLIGLAWGDAGGYKAADMWGPSSDPAWERNDPTQIPL 240
Db 179 DQFIYAGSLISALMDSSQGIPEQLIGLAWGDAGGYKAADMWGPSPNDPAWRNDPILQAGKL 238

QY 241 VANNTLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDQDAYNAGGHNAVFNFPNG 300
Db 239 VANNTLWVYCGNGTPELGGTNPVPAEFLNFVHGSNLKFDQDAYNAGGHNAVFNLNADG 298

QY 301 THSWEYWGALNMAKGDLOSSLGA 324
Db 299 THSWEYWGALNMAKDPDLQNTLMA 322
```

```

; TELEFAX: 206.269.0563
; INFORMATION FOR SEQ ID NO: 31:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-705-347A-31

```

```

Query Match      83.3%; Score 1454; DB 3; Length 327;
Best Local Similarity 83.0%; Pred. No. 4.5e-133;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;

```

```

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MIDVSGKIRAWGRLLVG--AAATPLSLISLAGGAATASAFSRPGLPVEYLQVPSMGR 58
QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
DB 59 TIKVFQFONGGNSPAVYLLDGLRAQDDYNGWDINTSAFEWYQSGLSIVMPVGGQSSFFS 118
QY 121 DWYSPACGKAGCCTYKWTETLTSELPOWLSANRAVKPTGSAAGLSMAGSSAMTILAAVHP 180
DB 119 DWYSPACGKAGCCTYKWTETLTSELPKWLSANRSVKSTGSAVGLSMAGSSALILAAVHP 178
QY 181 QQFIYAGSLALDPSQGMGSPSLIGLAMGDAGGYKAAADMWGPSSDPAWERNDPQQIPKL 240
DB 179 DQFIYAGSLALMDSSQGIPEQLIGLAMGDAGGYKAAADMWGPSSDPAWQRNDPILQAGKL 238
QY 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFQDAYNAAGGHNAVENFPNG 300
DB 239 VANNTHLWVYCGNGTPELGGTNPVPAEFLNFVHGSNLKFQDAYNGAGGHNAVENLNADG 298
QY 301 THSWEYWGALNNAKMDLQSSLSGA 324
DB 299 THSWEYWGALNNAKMDLQNTLMA 322

```

Search completed: February 5, 2004, 17:41:23
Job time : 23.6667 secs

```

RESULT 30
US-09-324-542-31
; Sequence 31, Application US/09324542
; Patent No. 6328978
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: PrestiGe, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c1
; CURRENT APPLICATION NUMBER: US/09/324,542
; EARLIER APPLICATION NUMBER: 1999-06-02
; EARLIER FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
; US-09-324-542-31

```

```

Query Match      83.3%; Score 1454; DB 4; Length 327;
Best Local Similarity 83.0%; Pred. No. 4.5e-133;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;

```

```

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MIDVSGKIRAWGRLLVG--AAATPLSLISLAGGAATASAFSRPGLPVEYLQVPSMGR 58
QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFFS 120
DB 59 TIKVFQFONGGNSPAVYLLDGLRAQDDYNGWDINTSAFEWYQSGLSIVMPVGGQSSFFS 118

```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:38:58 ; Search time 44.881 Seconds
(without alignments)
1516.217 Million cell updates/sec

Title: US-09-805-427A-2
Perfect score: 1745
Sequence: 1 MTDSVRKIRAWGRRLMIGTA.....YWGQLNAMKGDLOSSLGAG 325

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 788294

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1745	100.0	325	10	US-09-805-427A-2
2	1745	100.0	325	11	US-09-880-505-33
3	1745	100.0	325	11	US-09-952-554-1
4	1745	100.0	325	12	US-09-872-505-2
5	1745	100.0	325	14	US-10-051-643-33
6	1706	97.8	323	11	US-09-880-505-35
7	1706	97.8	323	14	US-10-051-643-35
8	1612	92.4	364	9	US-09-051-755-16
9	1612	92.4	364	15	US-10-231-055-16
10	1568.5	89.9	1016	12	US-10-369-983-18
11	1557	89.2	1010	12	US-10-369-983-4
12	1550	88.8	403	9	US-09-791-171-173
13	1550	88.8	403	10	US-09-805-427A-4
14	1550	88.8	403	12	US-09-804-980-173
15	1550	88.8	403	12	US-09-872-505-4
					Sequence 2, Appli
					Sequence 33, Appl
					Sequence 1, Appli
					Sequence 2, Appli
					Sequence 33, Appl
					Sequence 35, Appl
					Sequence 16, Appl
					Sequence 18, Appl
					Sequence 4, Appli
					Sequence 173, App
					Sequence 4, Appli
					Sequence 173, App
					Sequence 4, Appli

16	1550	88.8	404	9	US-09-791-171-172
17	1550	88.8	404	10	US-09-805-427A-3
18	1550	88.8	404	12	US-09-804-980-172
19	1550	88.8	404	12	US-09-872-505-3
20	1522	87.2	330	11	US-09-952-554-2
21	1454	83.3	327	11	US-09-880-505-31
22	1454	83.3	327	14	US-10-051-643-31
23	1407	80.6	338	11	US-09-880-505-32
24	1407	80.6	338	11	US-09-880-505-34
25	1407	80.6	338	14	US-10-051-643-32
26	1407	80.6	338	14	US-10-051-643-34
27	1339.5	76.8	364	9	US-09-051-755-17
28	1339.5	76.8	364	15	US-10-231-055-17
29	1335.5	76.5	330	11	US-09-880-505-30
30	1335.5	76.5	330	14	US-10-051-643-30
31	1195	68.5	340	11	US-09-880-505-37
32	1195	68.5	340	14	US-10-051-643-37
33	1179	67.6	326	11	US-09-880-505-43
34	1179	67.6	326	12	US-10-205-979-29
35	1179	67.6	326	14	US-10-051-643-43
36	1170	67.0	333	11	US-09-880-505-36
37	1170	67.0	333	14	US-10-051-643-36
38	1151.5	66.0	334	11	US-09-880-505-45
39	1151.5	66.0	334	12	US-10-205-979-31
40	1151.5	66.0	334	14	US-10-051-643-45
41	1093.5	62.7	364	9	US-09-051-755-14
42	1093.5	62.7	364	15	US-10-231-055-14
43	1068.5	61.2	364	9	US-09-051-755-15
44	1068.5	61.2	364	15	US-10-231-055-15
45	718	41.1	364	9	US-09-051-755-18
46	718	41.1	364	15	US-10-231-055-18
47	644	36.9	161	11	US-09-880-505-44
48	644	36.9	161	14	US-10-051-643-44
49	637	36.5	161	11	US-09-880-505-70
50	637	36.5	161	12	US-10-205-979-30
51	637	36.5	161	14	US-10-051-643-70
52	636	36.4	161	11	US-09-880-505-158
53	636	36.4	161	12	US-10-205-979-32
54	636	36.4	161	14	US-10-051-643-158
55	519	29.7	299	9	US-09-791-171-42
56	502.5	28.8	298	12	US-09-804-980-42
57	470	26.9	657	10	US-09-738-626-6670
58	415	23.8	341	10	US-09-738-626-6673
59	303.5	17.4	411	10	US-09-738-626-4649
60	279	16.0	462	10	US-09-738-626-4528
61	277	15.9	483	10	US-09-738-626-5891
62	262	15.0	626	10	US-09-948-722-2
63	238.5	13.7	365	10	US-09-738-626-3882
64	209	12.0	40	10	US-09-953-510-15
65	209	12.0	40	10	US-09-953-510-16
66	209	12.0	40	10	US-09-953-510-27
67	209	12.0	40	10	US-09-953-457-13
68	209	12.0	40	12	US-10-147-255-15
69	209	12.0	40	12	US-10-147-255-16
70	209	12.0	40	12	US-10-147-255-27
71	209	12.0	40	12	US-09-953-413-15
72	209	12.0	40	12	US-09-953-413-16
73	209	12.0	40	12	US-09-953-413-27
74	202	11.6	40	10	US-09-953-510-28
75	202	11.6	40	10	US-10-147-255-28
76	202	11.6	40	12	US-09-953-413-28
77	119	6.8	60	10	US-09-953-510-26
78	119	6.8	60	12	US-10-147-255-26
79	119	6.8	60	12	US-09-953-413-26
80	112	6.4	60	10	US-09-953-510-25
81	112	6.4	60	12	US-10-147-255-25
82	112	6.4	60	12	US-09-953-413-25
83	111.5	6.4	106	15	US-09-864-408A-1538
84	111.5	6.4	376	15	US-10-156-761-12208
85	109	6.2	635	11	US-09-851-847-5
86	108	6.2	22	10	US-09-953-510-17
87	108	6.2	22	12	US-10-147-255-17
88	108	6.2	22	12	US-09-953-413-17

89 106.5 6.1 781 15 US-10-156-761-12689 Sequence 12689, A
90 105 6.0 20 10 US-09-813-333-65 Sequence 65, Appl

ALIGNMENTS

```
RESULT 1
US-09-805-427A-2
; Sequence 2, Application US/09805427A
; Patent No. US20020176867A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Hybrids of M. tuberculosis Antigens
; FILE REFERENCE: 670001-2002.5
; CURRENT APPLICATION NUMBER: US/09/805,427A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(40)
US-09-805-427A-2

Query Match      100.0%; Score 1745; DB 10; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.1e-157;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIIVMPVGGSSFFYS 120
DB 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIIVMPVGGSSFFYS 120
QY 121 DWSPACGKAGCQTKWETFLTSELPQWLSANRAVKPTGSAAGLSMAGSSAMILLAAHP 180
DB 121 DWSPACGKAGCQTKWETFLTSELPQWLSANRAVKPTGSAAGLSMAGSSAMILLAAHP 180
QY 181 QOFTYAGSLSALLDPSQGMPSLIGLMDAGGYKAADMWGPSSDPAPERNDPTQOI PKL 240
DB 181 QOFTYAGSLSALLDPSQGMPSLIGLMDAGGYKAADMWGPSSDPAPERNDPTQOI PKL 240
QY 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDQAYNAAGHNAVFNPNG 300
DB 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDQAYNAAGHNAVFNPNG 300
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
DB 301 THSWEYGAQLNAMKGDLOSSLGAG 325
```

```
RESULT 2
US-09-880-505-33
; Sequence 33, Application US/09880505
; Publication No. US2003000796A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; FILE REFERENCE: 11000.1007c2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
```

```
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-880-505-33

Query Match      100.0%; Score 1745; DB 11; Length 325;
Best Local Similarity 100.0%; Pred. No. 6.1e-157;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIIVMPVGGSSFFYS 120
DB 61 DIKVFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFEWYQSGLSIIVMPVGGSSFFYS 120
QY 121 DWSPACGKAGCQTKWETFLTSELPQWLSANRAVKPTGSAAGLSMAGSSAMILLAAHP 180
DB 121 DWSPACGKAGCQTKWETFLTSELPQWLSANRAVKPTGSAAGLSMAGSSAMILLAAHP 180
QY 181 QOFTYAGSLSALLDPSQGMPSLIGLMDAGGYKAADMWGPSSDPAPERNDPTQOI PKL 240
DB 181 QOFTYAGSLSALLDPSQGMPSLIGLMDAGGYKAADMWGPSSDPAPERNDPTQOI PKL 240
QY 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDQAYNAAGHNAVFNPNG 300
DB 241 VANNTRLWVYCGNGTPELGGANIPAEFLNFVRSSNLKFDQAYNAAGHNAVFNPNG 300
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
DB 301 THSWEYGAQLNAMKGDLOSSLGAG 325

RESULT 3
US-09-952-554-1
; Sequence 1, Application US/09952554
; Publication No. US20030036104A1
; GENERAL INFORMATION:
; APPLICANT: Wallis, Robert S.
; TITLE OF INVENTION: DETECTION OF MYCOBACTERIA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/952,554
; FILING DATE: 14-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,347
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: CASE-02270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 1:
```



```
SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030036104A1 Relevant
TOPOLOGY: No. US20030036104A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-952-554-1

Query Match
Best Local Similarity 100.0%; Score 1745; DB 11; Length 325;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYS 120
DB 61 DIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYS 120
QY 121 DWYSPACGKAGCQTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
DB 121 DWYSPACGKAGCQTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
QY 181 QOFIYAGSLALLDPSCQMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNDDPTQIIPKL 240
DB 181 QOFIYAGSLALLDPSCQMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNDDPTQIIPKL 240
QY 241 VANNTLWYCCNGTNPNEELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPPNG 300
DB 241 VANNTLWYCCNGTNPNEELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPPNG 300
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
DB 301 THSWEYGAQLNAMKGDLOSSLGAG 325

RESULT 4
US-09-872-505-2
; Sequence 2, Application US/09872505
; Publication No. US20040013685A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Nucleic Acid Fragments Derived From M. Tuberculosis
; FILE REFERENCE: 670001-2002.6
; CURRENT APPLICATION NUMBER: US/09/872,505
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: (1)..(40)
US-09-872-505-2

Query Match
Best Local Similarity 100.0%; Score 1745; DB 12; Length 325;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYS 120
DB 61 DIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYS 120
QY 121 DWYSPACGKAGCQTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
DB 121 DWYSPACGKAGCQTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180

SEQUENCE CHARACTERISTICS:
LENGTH: 325 amino acids
TYPE: amino acid
STRANDEDNESS: No. US20030036104A1 Relevant
TOPOLOGY: No. US20030036104A1 Relevant
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-952-554-1

Query Match
Best Local Similarity 100.0%; Score 1745; DB 11; Length 325;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYS 120
DB 61 DIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYS 120
QY 121 DWYSPACGKAGCQTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
DB 121 DWYSPACGKAGCQTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
QY 181 QOFIYAGSLALLDPSCQMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNDDPTQIIPKL 240
DB 181 QOFIYAGSLALLDPSCQMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNDDPTQIIPKL 240
QY 241 VANNTLWYCCNGTNPNEELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPPNG 300
DB 241 VANNTLWYCCNGTNPNEELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPPNG 300
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
DB 301 THSWEYGAQLNAMKGDLOSSLGAG 325

RESULT 5
US-10-051-643-33
; Sequence 33, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium vaccae
; FILE REFERENCE: 11000.1008C2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 33
; LENGTH: 325
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-051-643-33

Query Match
Best Local Similarity 100.0%; Score 1745; DB 14; Length 325;
Matches 325; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLVPLGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYS 120
DB 61 DIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGQSSFYS 120
QY 121 DWYSPACGKAGCQTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
DB 121 DWYSPACGKAGCQTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHP 180
QY 181 QOFIYAGSLALLDPSCQMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNDDPTQIIPKL 240
DB 181 QOFIYAGSLALLDPSCQMGPSLIGLAMGDAGGYKAADMWGPSSDPAWERNDDPTQIIPKL 240
QY 241 VANNTLWYCCNGTNPNEELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPPNG 300
DB 241 VANNTLWYCCNGTNPNEELGGANI PAEFLNFVRSSNLKFQDAYNAAGHNAVFNPPNG 300
QY 301 THSWEYGAQLNAMKGDLOSSLGAG 325
DB 301 THSWEYGAQLNAMKGDLOSSLGAG 325

RESULT 6
US-09-880-505-35
; Sequence 35, Application US/09880505
; Publication No. US20030007976A1
```

```

; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Mycobacterium bovis
US-09-880-505-35

Query Match          97.8%; Score 1706; DB 11; Length 323;
Best Local Similarity 98.5%; Pred. No. 3e-153;
Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIWMVPGQSSPYS 120
DB 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIWMVPGQSSPYS 120
QY 121 DWSPACGKACQYKWTETLITSELPWLSANRAVKPTGSAAGLSWAGSSAMILAAHP 180
DB 121 DWSPACGKACQYKWTETLITSELPWLSANRAVKPTGSAAGLSWAGSSAMILAAHP 180
QY 181 QOFIYAGLSALLDPSQGMGSLIGLAMGDAGGYKAADMWGPSSDPAWERNDTQQIPKL 240
DB 181 QOFIYAGLSALLDPSQGMG--LIGLAMGDAGGYKAADMWGPSSDPAWERNDTQQIPKL 238
QY 241 VANNTRLWVYCGNCTPNELGGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVENFPNG 300
DB 239 VANNTRLWVYCGNCTPNELGGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVENFPNG 298
QY 301 THSWEYGAQINAMKGDLOSSLGAG 325
DB 299 THSWEYGAQINAMKGDLOSSLGAG 323

RESULT 7
US-10-051-643-35
; Sequence 35, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium Vaccae
; FILE REFERENCE: 11000.1008c2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 35
; LENGTH: 323
; TYPE: PRT
; ORGANISM: Mycobacterium bovis

```

```

US-10-051-643-35

Query Match          97.8%; Score 1706; DB 14; Length 323;
Best Local Similarity 98.5%; Pred. No. 3e-153;
Matches 320; Conservative 0; Mismatches 3; Indels 2; Gaps 1;

QY 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
DB 1 MTDVSRKIRAWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
QY 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIWMVPGQSSPYS 120
DB 61 DIKVFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIWMVPGQSSPYS 120
QY 121 DWSPACGKACQYKWTETLITSELPWLSANRAVKPTGSAAGLSWAGSSAMILAAHP 180
DB 121 DWSPACGKACQYKWTETLITSELPWLSANRAVKPTGSAAGLSWAGSSAMILAAHP 180
QY 181 QOFIYAGLSALLDPSQGMGSLIGLAMGDAGGYKAADMWGPSSDPAWERNDTQQIPKL 240
DB 181 QOFIYAGLSALLDPSQGMG--LIGLAMGDAGGYKAADMWGPSSDPAWERNDTQQIPKL 238
QY 241 VANNTRLWVYCGNCTPNELGGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVENFPNG 300
DB 239 VANNTRLWVYCGNCTPNELGGANIPAEFLENFVRSSNLKFQDAYNAAGHNAVENFPNG 298
QY 301 THSWEYGAQINAMKGDLOSSLGAG 325
DB 299 THSWEYGAQINAMKGDLOSSLGAG 323

RESULT 8
US-09-051-755-16
; Sequence 16, Application US/09051755
; Patent No. US20010018048A1
; GENERAL INFORMATION:
; APPLICANT: LEER, Robert J
; APPLICANT: POWELS, Pieter H
; APPLICANT: CONWAY, Patricia L
; TITLE OF INVENTION: NOVEL ADHERENCE FACTORS OF NON PATHOGENIC
; TITLE OF INVENTION: MICROORGANISMS AND APPLICATIONS THEREOF FOR SCREENING
; TITLE OF INVENTION: MICROORGANISMS FOR SPECIFIC PROBIOTIC PROPERTIES; NOVEL
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS AND FOOD ADDITIVES C
; FILE REFERENCE: LEER et al. 09/051,755
; CURRENT APPLICATION NUMBER: US/09/051,755
; CURRENT FILING DATE: 1998-09-30
; EARLIER APPLICATION NUMBER: PCT/NL95/00367
; EARLIER FILING DATE: 1995-10-20
; EARLIER APPLICATION NUMBER: PCT/NL96/00409
; EARLIER FILING DATE: 1996-10-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 16
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: aligned aa
; OTHER INFORMATION: sequence of Mycobacterium Mtu85b
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)...(364)
; OTHER INFORMATION: Xaa is any amino acid.
US-09-051-755-16

Query Match          92.4%; Score 1612; DB 9; Length 364;
Best Local Similarity 92.6%; Pred. No. 2.9e-144;
Matches 325; Conservative 0; Mismatches 0; Indels 26; Gaps 12;

QY 1 MTDVSRKIRAAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPS 55
DB 3 MTDVSRKIRAXXXXXXWGRRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPS 62

```

```

QY 56 P-SMGRDIKV-QFQSGGNSPAVLLDGLRAQDDY--NGWDINTPAFEWYQSG-LSIVM 110
Db 63 PASMGRDIKVQFQSGGNSPAVLLDGLRAQDDYXXNGWDINTPAFEWYQSGXLSIVM 122
QY 111 PVGGQSSFYSDWYSPACGKAGCQTYKWET-FLTSELQWLSANRAVKPTGSAAGL-----165
Db 123 PVGGQSSFYSDWYSPACGKAGCQTYKWETXFLTSELQWLSANRAVKPTGSAAGLXXXX 182
QY 166 --SMAGSS---AMILAAHYHQPIFI---YAGSLALLDPQSGMGPSSLIGLAMDAGGYKAA 217
Db 183 XXSMAGSSXXXAMILAAHYHQPIFIXXXAGSLALLDPQSGMGPSSLIGLAMDAGGYKAA 242
QY 218 DMWGSSDDPAWERNDDTQOIPKLVANNTRLWVYCGNGTNEELGGANI PAEFLNFVRSSN 277
Db 243 DMWGSSDDPAWERNDDTQOIPKLVANNTRLWVYCGNGTNEELGGANI PAEFLNFVRSSN 302
QY 278 LKFDQDAYNAAGGHNAVFNPFPNGTHSWE--YMGALNAMKGDLOSSL-GAG 325
Db 303 LKFDQDAYNAAGGHNAVFNPFPNGTHSWEXXYMGALNAMKGDLOSSLXGAG 353
RESULT 9
US-10-231-055-16
; Sequence 16, Application US/10231055
; Publication No. US2003069578A1
; GENERAL INFORMATION:
; APPLICANT: LEER, Robert J
; APPLICANT: FOWELLS, Pieter H
; APPLICANT: CONWAY, Patricia L
; TITLE OF INVENTION: NOVEL ADHERENCE FACTORS OF NON PATHOGENIC
; TITLE OF INVENTION: MICROORGANISMS AND APPLICATIONS THEREOF FOR SCREENING
; TITLE OF INVENTION: MICROORGANISMS FOR SPECIFIC PROBIOTIC PROPERTIES; NOVEL
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS AND FOOD ADDITIVES C
; FILE REFERENCE: LEER et al. 09/051,755
; CURRENT APPLICATION NUMBER: US/10/231,055
; PRIOR FILING DATE: 2002-08-30
; PRIOR APPLICATION NUMBER: US/09/051,755
; PRIOR FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/NL95/00367
; PRIOR FILING DATE: 1995-10-20
; PRIOR APPLICATION NUMBER: PCT/NL96/00409
; PRIOR FILING DATE: 1996-10-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 16
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:aligned aa
; OTHER INFORMATION: sequence of Mycobacterium Mtu85b
; NAME/KEY: UNSURE
; LOCATION: (1)..(364)
; OTHER INFORMATION: Xaa is any amino acid.
US-10-231-055-16
Query Match 92.4%; Score 1612; DB 15; Length 364;
Best Local Similarity 92.6%; Pred. No. 2.9e-144;
Matches 325; Conservative 0; Mismatches 0; Indels 26; Gaps 12;
QY 1 MTDVSRKIRA-----WGRRLMIGTAAAVLPGLVLGAGGAATAGASRPG-LPVEYLQVPS 55
Db 3 MTDVSRKIRAXXXXWGRRLMIGTAAAVLPGLVLGAGGAATAGASRPG-LPVEYLQVPS 62
QY 56 P-SMGRDIKV-QFQSGGNSPAVLLDGLRAQDDY--NGWDINTPAFEWYQSG-LSIVM 110
Db 63 PASMGRDIKVQFQSGGNSPAVLLDGLRAQDDYXXNGWDINTPAFEWYQSGXLSIVM 122
QY 111 PVGGQSSFYSDWYSPACGKAGCQTYKWET-FLTSELQWLSANRAVKPTGSAAGL-----165
Db 123 PVGGQSSFYSDWYSPACGKAGCQTYKWETXFLTSELQWLSANRAVKPTGSAAGLXXXX 182

```

```

QY 166 --SMAGSS---AMILAAHYHQPIFI---YAGSLALLDPQSGMGPSSLIGLAMDAGGYKAA 217
Db 183 XXSMAGSSXXXAMILAAHYHQPIFIXXXAGSLALLDPQSGMGPSSLIGLAMDAGGYKAA 242
QY 218 DMWGSSDDPAWERNDDTQOIPKLVANNTRLWVYCGNGTNEELGGANI PAEFLNFVRSSN 277
Db 243 DMWGSSDDPAWERNDDTQOIPKLVANNTRLWVYCGNGTNEELGGANI PAEFLNFVRSSN 302
QY 278 LKFDQDAYNAAGGHNAVFNPFPNGTHSWE--YMGALNAMKGDLOSSL-GAG 325
Db 303 LKFDQDAYNAAGGHNAVFNPFPNGTHSWEXXYMGALNAMKGDLOSSLXGAG 353
RESULT 10
US-10-369-983-18
; Sequence 18, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 18
; LENGTH: 1016
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:fusion protein
; OTHER INFORMATION: MTB103F (MTB72F-85b)
US-10-369-983-18
Query Match 89.9%; Score 1568.5; DB 12; Length 1016;
Best Local Similarity 94.5%; Pred. No. 1.5e-139;
Matches 292; Conservative 4; Mismatches 12; Indels 1; Gaps 1;
QY 18 GTAAAVLPLGLVGLAG-GAATAGAFSRPGLPVEYLQVPSMGRDIKVQFQSGGNSPAV 76
Db 708 GDSGGPVVNGLQVVGVMNTAASGTSRFPCLPVEYLQVPSMGRDIKVQFQSGGNSPAV 767
QY 77 YLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSPACGKAGCQTYK 136
Db 768 YLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSPACGKAGCQTYK 827
QY 137 WETFLTSELQWLSANRAVKPTGSAAGLSWAGSSAMILAAHYHQPIFIYAGSLALLDPS 196
Db 828 WETFLTSELQWLSANRAVKPTGSAAGLSWAGSSAMILAAHYHQPIFIYAGSLALLDPS 887
QY 197 QCMGPSLLIGLAMDAGGYKAAADMWGPSSDDPAWERNDDTQOIPKLVANNTRLWVYCGNGTP 256
Db 888 QCMGPSLLIGLAMDAGGYKAAADMWGPSSDDPAWERNDDTQOIPKLVANNTRLWVYCGNGTP 947
QY 257 NELGGANI PAEFLNFVRSSNLKFODAYNAAGGHNAVFNPFPNGTHSWEYGAQLNAMKG 316
Db 948 NELGGANI PAEFLNFVRSSNLKFODAYNAAGGHNAVFNPFPNGTHSWEYGAQLNAMKG 1007
QY 317 DLQSSLGAG 325
Db 1008 DLQSSLGAG 1016

```

```

RESULT 11
US-10-369-983-4
; Sequence 4, Application US/10369983
; Publication No. US20030235593A1
; GENERAL INFORMATION:

```

```
; APPLICANT: Skeiky, Yasir
; APPLICANT: Guderian, Jeff
; APPLICANT: Reed, Steven
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Fusion Proteins of Mycobacterium Tuberculosis
; FILE REFERENCE: 014058-009081US
; CURRENT APPLICATION NUMBER: US/10/369,983
; CURRENT FILING DATE: 2003-02-18
; PRIOR APPLICATION NUMBER: US 60/357,351
; PRIOR FILING DATE: 2002-02-15
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 1010
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:MTB-102F fusion
; OTHER INFORMATION: protein
US-10-369-983-4

Query Match      89.2%; Score 1557; DB 12; Length 1010;
Best Local Similarity 82.9%; Pred. No. 1.8e-138;
Matches 296; Conservative 3; Mismatches 18; Indels 40; Gaps 2;

QY 9 RAWGRRLLMIGTAAAVLPL-----GLVGLAGGAATAGA----- 40
DB 654 QAWAANQAVTPARALPLTSLTSAERGPQMLGGLPVQMGARAGGSLGVLRVPRP 713

QY 41 -----FSRGLPVEYLQVPSMGRDIKVQFQSGGNNSPAVYLLDGLRAQDDY 88
DB 714 YVMPHSPAAGKLSRGLPVEYLQVPSMGRDIKVQFQSGGNNSPAVYLLDGLRAQDDY 773

QY 89 NGHDINTPAPEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCQTYKWETFTLSELPOW 148
DB 774 NGWDINTPAPEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCQTYKWETFTLSELPOW 833

QY 149 LSNARVKTGSAAGLSMAGSSAMILAAVHPQFIYAGSLSALLDPSQGMGSLIGLAM 208
DB 834 LSNARVKTGSAAGLSMAGSSAMILAAVHPQFIYAGSLSALLDPSQGMGSLIGLAM 893

QY 209 GDAGGYKAADWGFSSDPAWERNPTQIQPKLVANTRLWYCGNGTPELGGANIPAEF 268
DB 894 GDAGGYKAADWGFSSDPAWERNPTQIQPKLVANTRLWYCGNGTPELGGANIPAEF 953

QY 269 LENFVRSNLKFDAYNAAGHNVAFFPPNGTHSWEYWGQALNAMKGDLOSSLGAG 325
DB 954 LENFVRSNLKFDAYNAAGHNVAFFPPNGTHSWEYWGQALNAMKGDLOSSLGAG 1010

RESULT 12
US-09-791-171-173
; Sequence 173, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
```

```
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 173
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-173

Query Match      88.8%; Score 1550; DB 9; Length 403;
Best Local Similarity 99.3%; Pred. No. 2.5e-138;
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 38 AGAFSRPGLPVEYLQVPSMGRDIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPA 97
DB 116 AKLFSPRGLPVEYLQVPSMGRDIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPA 175

QY 98 FEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCQTYKWETFTLSELPOWLSANRAVKP 157
DB 176 FEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCQTYKWETFTLSELPOWLSANRAVKP 235

QY 158 TGSAAIGLSMAGSSAMILAAVHPQFIYAGSLSALLDPSQGMGSLIGLAMGDAGGYKAA 217
DB 236 TGSAAIGLSMAGSSAMILAAVHPQFIYAGSLSALLDPSQGMGSLIGLAMGDAGGYKAA 295

QY 218 DMGPPSDPAWERNPTQIQPKLVANTRLWYCGNGTPELGGANIPAEFLENFVRSN 277
DB 296 DMGPPSDPAWERNPTQIQPKLVANTRLWYCGNGTPELGGANIPAEFLENFVRSN 355

QY 278 LKFDQAYNAAGHNVAFFPPNGTHSWEYWGQALNAMKGDLOSSLGAG 325
DB 356 LKFDQAYNAAGHNVAFFPPNGTHSWEYWGQALNAMKGDLOSSLGAG 403

RESULT 13
US-09-805-427a-4
; Sequence 4, Application US/09805427A
; Patent No. US20020176867A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Hybrids of M. tuberculosis Antigens
; FILE REFERENCE: 670001-2002.5
; CURRENT APPLICATION NUMBER: US/09/805,427A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant Fusion protein ESAT-6-Ag85B
US-09-805-427a-4

Query Match      88.8%; Score 1550; DB 10; Length 403;
Best Local Similarity 99.3%; Pred. No. 2.5e-138;
Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 38 AGAFSRPGLPVEYLQVPSMGRDIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPA 97
DB 116 AKLFSPRGLPVEYLQVPSMGRDIKVQFQSGGNNSPAVYLLDGLRAQDDYNGWDINTPA 175

QY 98 FEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCQTYKWETFTLSELPOWLSANRAVKP 157
DB 176 FEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCQTYKWETFTLSELPOWLSANRAVKP 235

QY 158 TGSAAIGLSMAGSSAMILAAVHPQFIYAGSLSALLDPSQGMGSLIGLAMGDAGGYKAA 217
DB 236 TGSAAIGLSMAGSSAMILAAVHPQFIYAGSLSALLDPSQGMGSLIGLAMGDAGGYKAA 295
```

QY 218 DMWGPSSDPAWERNPTQOIPKLVANNTRLVVYCGNGTNELGANIPAEFFLENFVRSSN 277
 Db 296 DMWGPSSDPAWERNPTQOIPKLVANNTRLVVYCGNGTNELGANIPAEFFLENFVRSSN 355
 QY 278 LKFQDAYNAAGGHNAVFPPNGTHSWYWGQALNWKDGLQSSLGAG 325
 Db 356 LKFQDAYNAAGGHNAVFPPNGTHSWYWGQALNWKDGLQSSLGAG 403
 RESULT 14
 US-09-804-980-173
 ; Sequence 173, Application US/09804980
 ; Publication No. US20030147897A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Statens Serum Institut
 ; APPLICANT: Anderson, Peter
 ; TITLE OF INVENTION: M. Tuberculosis Antigens
 ; FILE REFERENCE: 670001-2002.4
 ; CURRENT APPLICATION NUMBER: US/09/804,980
 ; CURRENT FILING DATE: 2001-03-12
 ; NUMBER OF SEQ ID NOS: 257
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 173
 ; LENGTH: 403
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 US-09-804-980-173

Query Match 88.8%; Score 1550; DB 12; Length 403;
 Best Local Similarity 99.3%; Pred. No. 2.5e-138;
 Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 38 AGAFSPGLPVEYLQVPSPSMGRIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPA 97
 Db 116 AKLFSPGLPVEYLQVPSPSMGRIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPA 175
 QY 98 FEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCCTYKWEFTLTSELPOWLSANRAVKP 157
 Db 176 FEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCCTYKWEFTLTSELPOWLSANRAVKP 235
 QY 158 TGSAAIGLSMAGSSAMILAAHYHPQOPIYAGSLSALLDPSQGMGPSLIGLAMGDAGYKAA 217
 Db 296 DMWGPSSDPAWERNPTQOIPKLVANNTRLVVYCGNGTNELGANIPAEFFLENFVRSSN 355
 QY 278 LKFQDAYNAAGGHNAVFPPNGTHSWYWGQALNWKDGLQSSLGAG 325
 Db 356 LKFQDAYNAAGGHNAVFPPNGTHSWYWGQALNWKDGLQSSLGAG 403

RESULT 15
 US-09-872-505-4
 ; Sequence 4, Application US/09872505
 ; Publication No. US20040013685A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Statens Serum Institut
 ; TITLE OF INVENTION: Nucleic Acid Fragments Derived From M. Tuberculosis
 ; FILE REFERENCE: 670001-2002.6
 ; CURRENT APPLICATION NUMBER: US/09/872,505
 ; CURRENT FILING DATE: 2001-06-01
 ; NUMBER OF SEQ ID NOS: 12
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 4
 ; LENGTH: 403
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Recombinant Fusion protein ESAT-6-Ag85B
 US-09-872-505-4

Query Match 88.8%; Score 1550; DB 12; Length 403;
 Best Local Similarity 99.3%; Pred. No. 2.5e-138;
 Matches 286; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 38 AGAFSPGLPVEYLQVPSPSMGRIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPA 97
 Db 116 AKLFSPGLPVEYLQVPSPSMGRIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPA 175
 QY 98 FEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCCTYKWEFTLTSELPOWLSANRAVKP 157
 Db 176 FEWYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCCTYKWEFTLTSELPOWLSANRAVKP 235
 QY 158 TGSAAIGLSMAGSSAMILAAHYHPQOPIYAGSLSALLDPSQGMGPSLIGLAMGDAGYKAA 217
 Db 296 DMWGPSSDPAWERNPTQOIPKLVANNTRLVVYCGNGTNELGANIPAEFFLENFVRSSN 355
 QY 278 LKFQDAYNAAGGHNAVFPPNGTHSWYWGQALNWKDGLQSSLGAG 325
 Db 356 LKFQDAYNAAGGHNAVFPPNGTHSWYWGQALNWKDGLQSSLGAG 403

RESULT 16
 US-09-791-171-172
 ; Sequence 172, Application US/09791171
 ; Patent No. US20020094336A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ANDERSEN, Peter
 ; APPLICANT: NIELSEN, Rikke
 ; APPLICANT: OETTINGER, Thomas
 ; APPLICANT: RASMUSSEN, Peter Birk
 ; APPLICANT: ROSENKRANDS, Ida
 ; APPLICANT: WELLDINGH, Karin
 ; APPLICANT: FLORIO, Walter
 ; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
 ; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
 ; FILE REFERENCE: 670001-2002.1
 ; CURRENT APPLICATION NUMBER: US/09/791,171
 ; CURRENT FILING DATE: 2001-02-20
 ; PRIOR APPLICATION NUMBER: 09/050,739
 ; PRIOR FILING DATE: 1998-03-30
 ; PRIOR APPLICATION NUMBER: 0376/97
 ; PRIOR FILING DATE: 1997-04-02
 ; PRIOR APPLICATION NUMBER: 1277/97
 ; PRIOR FILING DATE: 1997-11-10
 ; PRIOR APPLICATION NUMBER: 60/044,624
 ; PRIOR FILING DATE: 1997-04-18
 ; PRIOR APPLICATION NUMBER: 60/070,488
 ; PRIOR FILING DATE: 1998-01-05
 ; NUMBER OF SEQ ID NOS: 173
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 172
 ; LENGTH: 404
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 US-09-791-171-172

Query Match 88.8%; Score 1550; DB 9; Length 404;
 Best Local Similarity 99.7%; Pred. No. 2.5e-138;
 Matches 285; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 40 AFSRPGLPVEYLQVPSPSMGRIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPAPE 99
 Db 21 SFSRPGLPVEYLQVPSPSMGRIKVFQSGGNNSPAVLLDGLRAODDYNWDINTPAPE 80
 QY 100 WYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCCTYKWEFTLTSELPOWLSANRAVKP 159
 Db 81 WYQSGLSIVMPVGGQSSFYSDWYSPACGKAGCCTYKWEFTLTSELPOWLSANRAVKP 140
 QY 160 SAAIGLSMAGSSAMILAAHYHPQOPIYAGSLSALLDPSQGMGPSLIGLAMGDAGYKAA 219

Db 141 SAAIGLSMAGSSAMILAAAYHPQQFIYAGSLSALLDPSQGMGPSLIGLWGDAGGYKAAD 200
Qy 220 WGPSSDPAWERNPTQIQPKLVANNTLWYCGNGTNEELGGANIPAEFFLENFVRSSNLK 279
Db 201 WGPSSDPAWERNPTQIQPKLVANNTLWYCGNGTNEELGGANIPAEFFLENFVRSSNLK 260
Qy 280 FQDAYNAAGGHNAVFNPNGTHSWEYWGAGLNAKMGDLQSSLGAG 325
Db 261 FQDAYNAAGGHNAVFNPNGTHSWEYWGAGLNAKMGDLQSSLGAG 306

RESULT 17

US-09-805-427A-3
; Sequence 3, Application US/09805427A
; Patent No. US20020176867A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Hybrids of M. tuberculosis Antigens
; FILE REFERENCE: 670001-2002.5
; CURRENT APPLICATION NUMBER: US/09/805,427A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant Fusion protein Ag85B-ESAT-6
US-09-805-427A-3

Query Match 88.8%; Score 1550; DB 10; Length 404;
Best Local Similarity 99.7%; Pred. No. 2.5e-138;
Matches 285; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 40 AFSRPGLEVEYLQVPSPSMGRDIKVQFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFE 99
Db 21 SFSRPGLEVEYLQVPSPSMGRDIKVQFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFE 80
Qy 100 WYQSGLSIYMPVGGQSSFYSDWYSPACGKAGCQTYKWETFLTSELQWLSANRAVKPTG 159
Db 81 WYQSGLSIYMPVGGQSSFYSDWYSPACGKAGCQTYKWETFLTSELQWLSANRAVKPTG 140
Qy 160 SAAIGLSMAGSSAMILAAAYHPQQFIYAGSLSALLDPSQGMGPSLIGLWGDAGGYKAAD 219
Db 141 SAAIGLSMAGSSAMILAAAYHPQQFIYAGSLSALLDPSQGMGPSLIGLWGDAGGYKAAD 200
Qy 220 WGPSSDPAWERNPTQIQPKLVANNTLWYCGNGTNEELGGANIPAEFFLENFVRSSNLK 279
Db 201 WGPSSDPAWERNPTQIQPKLVANNTLWYCGNGTNEELGGANIPAEFFLENFVRSSNLK 260
Qy 280 FQDAYNAAGGHNAVFNPNGTHSWEYWGAGLNAKMGDLQSSLGAG 325
Db 261 FQDAYNAAGGHNAVFNPNGTHSWEYWGAGLNAKMGDLQSSLGAG 306

RESULT 18

US-09-804-980-172
; Sequence 172, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 172
; LENGTH: 404
; TYPE: PRT

; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-172

Query Match 88.8%; Score 1550; DB 12; Length 404;
Best Local Similarity 99.7%; Pred. No. 2.5e-138;
Matches 285; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 40 AFSRPGLEVEYLQVPSPSMGRDIKVQFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFE 99
Db 21 SFSRPGLEVEYLQVPSPSMGRDIKVQFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFE 80
Qy 100 WYQSGLSIYMPVGGQSSFYSDWYSPACGKAGCQTYKWETFLTSELQWLSANRAVKPTG 159
Db 81 WYQSGLSIYMPVGGQSSFYSDWYSPACGKAGCQTYKWETFLTSELQWLSANRAVKPTG 140
Qy 160 SAAIGLSMAGSSAMILAAAYHPQQFIYAGSLSALLDPSQGMGPSLIGLWGDAGGYKAAD 219
Db 141 SAAIGLSMAGSSAMILAAAYHPQQFIYAGSLSALLDPSQGMGPSLIGLWGDAGGYKAAD 200
Qy 220 WGPSSDPAWERNPTQIQPKLVANNTLWYCGNGTNEELGGANIPAEFFLENFVRSSNLK 279
Db 201 WGPSSDPAWERNPTQIQPKLVANNTLWYCGNGTNEELGGANIPAEFFLENFVRSSNLK 260
Qy 280 FQDAYNAAGGHNAVFNPNGTHSWEYWGAGLNAKMGDLQSSLGAG 325
Db 261 FQDAYNAAGGHNAVFNPNGTHSWEYWGAGLNAKMGDLQSSLGAG 306

RESULT 19

US-09-872-505-3
; Sequence 3, Application US/09872505
; Publication No. US20040013685A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Nucleic Acid Fragments Derived From M. Tuberculosis
; FILE REFERENCE: 670001-2002.6
; CURRENT APPLICATION NUMBER: US/09/872,505
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant Fusion protein Ag85B-ESAT-6
US-09-872-505-3

Query Match 88.8%; Score 1550; DB 12; Length 404;
Best Local Similarity 99.7%; Pred. No. 2.5e-138;
Matches 285; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 40 AFSRPGLEVEYLQVPSPSMGRDIKVQFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFE 99
Db 21 SFSRPGLEVEYLQVPSPSMGRDIKVQFQSGGNNSPAVLLDGLRAQDDYNGWDINTPAFE 80
Qy 100 WYQSGLSIYMPVGGQSSFYSDWYSPACGKAGCQTYKWETFLTSELQWLSANRAVKPTG 159
Db 81 WYQSGLSIYMPVGGQSSFYSDWYSPACGKAGCQTYKWETFLTSELQWLSANRAVKPTG 140
Qy 160 SAAIGLSMAGSSAMILAAAYHPQQFIYAGSLSALLDPSQGMGPSLIGLWGDAGGYKAAD 219
Db 141 SAAIGLSMAGSSAMILAAAYHPQQFIYAGSLSALLDPSQGMGPSLIGLWGDAGGYKAAD 200
Qy 220 WGPSSDPAWERNPTQIQPKLVANNTLWYCGNGTNEELGGANIPAEFFLENFVRSSNLK 279
Db 201 WGPSSDPAWERNPTQIQPKLVANNTLWYCGNGTNEELGGANIPAEFFLENFVRSSNLK 260
Qy 280 FQDAYNAAGGHNAVFNPNGTHSWEYWGAGLNAKMGDLQSSLGAG 325
Db 261 FQDAYNAAGGHNAVFNPNGTHSWEYWGAGLNAKMGDLQSSLGAG 306

RESULT 20

US-09-952-554-2
; Sequence 2, Application US/09952554
; Publication No. US20030036104A1
; GENERAL INFORMATION:
; APPLICANT: Wallis, Robert S.
; TITLE OF INVENTION: DETECTION OF MYCOBACTERIA
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MEDLEN & CARROLL, LLP
; STREET: 220 Montgomery Street, Suite 2200
; CITY: San Francisco
; STATE: California
; COUNTRY: United States of America
; ZIP: 94104
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/952,554
; FILING DATE: 14-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/690,347
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: MacKnight, Kamrin T.
; REGISTRATION NUMBER: 38,230
; REFERENCE/DOCKET NUMBER: CASE-02270
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 705-8410
; TELEFAX: (415) 397-8338
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 330 amino acids
; TYPE: amino acid
; STRANDEDNESS: No. US20030036104A1 Relevant
; TOPOLOGY: No. US20030036104A1 Relevant
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

Query Match 87.2%; Score 1522; DB 11; Length 330;

Best Local Similarity 84.9%; Pred. No. 8.4e-136;

Matches 275; Conservative 24; Mismatches 25; Indels 0; Gaps 0;

QY	1	MTDVSERKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSPGIPVEYLQVPSMGR	60
Db	1	MTDLSEKVRWGRRLVGLAAAVTLPLGLVGLAGGAATAGAFSPGIPVEYLQVPSMGR	60
QY	61	DIKVFQSGGNSPAVYLLDGLRAODDYGWDINTPAFEWYQSGLSIVMPVGGQSSFY	120
Db	61	DIKVFQSGGNSPAVYLLDGLRAODDYGWDINTPAFEWYQSGLSIVMPVGGQSSFY	120
QY	121	DWSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAATGLSMAGSSAMILAA	180
Db	121	DWYQACGKAGCSTYKWTFTLSELPQWLSANRAVKPTGSAATGLSMAGSSAMILAA	180
QY	181	QOPIYAGSLSALLDPGQMGPSLIGLAMDAGGYKAAADWGWSSDDPAWERNPTQI	240
Db	181	DOFIYAGSLSALLDPGQMGPSLIGLAMDAGGYKAAADWGWSSDDPAWERNPTQI	240
QY	241	VANNTLWVYCGNGTNEIAGGANIPAEFLENFVRSSNLKFDQAYNAGGHN	300
Db	241	VHNTLWVYCGNGTNEIAGGANIPAEFLENFVRSSNLKFDQAYNAGGHN	300
QY	301	THSWEYGAQLNAKMGDLOSSLGA	324
Db	301	THSWEYGAQLNAKMGDLOSSLGA	324

RESULT 21

US-09-880-505-31
; Sequence 31, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.11007c2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US/09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US/08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 31
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
; US-09-880-505-31

Query Match 83.3%; Score 1454; DB 11; Length 327;

Best Local Similarity 83.0%; Pred. No. 2.3e-129;

Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;

QY	1	MTDVSERKIRAWGRRLMIGTAAAVLPLGLVGLAGGAATAGAFSPGIPVEYLQVPSMGR	60
Db	1	MIDVSGKIRAWGRWLLVG--AAATLPSLISLAGGAATAGAFSPGIPVEYLQVPSMGR	58
QY	61	DIKVFQSGGNSPAVYLLDGLRAODDYGWDINTPAFEWYQSGLSIVMPVGGQSSFY	120
Db	59	TIKVFQSGGNSPAVYLLDGLRAODDYGWDINTPAFEWYQSGLSIVMPVGGQSSFY	118
QY	121	DWSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAATGLSMAGSSAMILAA	180
Db	119	DWSPACGKAGCQTYKWTFTLSELPQWLSANRAVKPTGSAATGLSMAGSSAMILAA	178
QY	181	QOPIYAGSLSALLDPGQMGPSLIGLAMDAGGYKAAADWGWSSDDPAWERNPTQI	240
Db	179	DOFIYAGSLSALLDPGQMGPSLIGLAMDAGGYKAAADWGWSSDDPAWERNPTQI	238
QY	241	VANNTLWVYCGNGTNEIAGGANIPAEFLENFVRSSNLKFDQAYNAGGHN	300
Db	239	VANNTLWVYCGNGTNEIAGGANIPAEFLENFVRSSNLKFDQAYNAGGHN	298
QY	301	THSWEYGAQLNAKMGDLOSSLGA	324
Db	299	THSWEYGAQLNAKMGDLOSSLGA	322

RESULT 22

US-10-051-643-31

; Sequence 31, Application US/10051643

; Publication No. US20020197265A1

; GENERAL INFORMATION:

; APPLICANT: Watson, James D.

; APPLICANT: Tan, Paul L. J.

; TITLE OF INVENTION: Methods and Compounds for the Treatment

; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory

; FILE REFERENCE: 11000.1008c2

; CURRENT APPLICATION NUMBER: US/10/051,643

; CURRENT FILING DATE: 2002-01-18

; PRIOR APPLICATION NUMBER: US09/156,181

; PRIOR FILING DATE: 1998-09-17

; PRIOR APPLICATION NUMBER: US 08/996,624

; PRIOR FILING DATE: 1997-12-23

; NUMBER OF SEQ ID NOS: 208

; SOFTWARE: FastSeq for Windows Version 3.0

```
; SEQ ID NO 31
; LENGTH: 327
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-051-643-31

Query Match      83.3%; Score 1454; DB 14; Length 327;
Best Local Similarity 83.0%; Pred. No. 2.3e-129;
Matches 269; Conservative 19; Mismatches 34; Indels 2; Gaps 1;

QY 1 MTIVSRKIRAWGRRLMTGTAARVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGR 60
Db 1 MIDVSGKIRAWGRLLVG--AAATPLSLISLAGGAATASAFSRPGLPVEYLQVPSMGR 58
QY 61 DIKVOFSGGNNSPAVYLLDGLRAQDDYNGWDINTPAFVWYQSGLSIVMPVGGSSPYS 120
Db 59 TIKVOFONGGNSPAVYLLDGLRAQDDYNGWDINTSAFVWYQSGLSIVMPVGGSSPYS 118
QY 121 DWSPACGKAGCQTYKWETFLTSELTPQLWLSANRAVKPTGSAAGLSMAGSSAMILAAHYH 180
Db 119 DWSPACGKAGCTTYKWETFLTSELTPKWL SANRSVKSTGSAVVGLSMAGSSALILAAHYH 178
QY 181 QOFYAGSLALLDPSQMGPSLIGLAMGAGGKAAADMGPPSDPAWERNDPQQIPKL 240
Db 179 DQFIYAGSLALMDSSQIEPQLTGLAMGAGGKAAADMGPPNDPAWQRNDPILQAGKL 238
QY 241 VANNTRLWVYCGNGTPELGCANIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNPFPNG 300
Db 239 VANNTHLWVYCGNGTPELGTNVPAEFLENFVHGSNLKFQDAYNGAGGHNAVFNLNADG 298
QY 301 THSWEYWGALNAMKGDLOSSLGA 324
Db 299 THSWEYWGALNAMKPDQLNTLAA 322

RESULT 23
US-09-880-505-32
; Sequence 32, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007c2
; CURRENT APPLICATION NUMBER: US/09/880,505
; PRIOR FILING DATE: 2001-06-13
; PRIOR FILING DATE: 1999-06-02
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-880-505-32

Query Match      80.6%; Score 1407; DB 11; Length 338;
Best Local Similarity 78.9%; Pred. No. 6.9e-125;
Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

QY 8 IRANGRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGRDQKVFQ 67
Db 11 VTGMSRLVVGAAVLVGLVAVGVTATAGAFSRPGLPVEYLQVPSMGRDQKVFQ 70
QY 68 SGGNNSPAVYLLDGLRAQDDYNGWDINTPAFVWYQSGLSIVMPVGGSSPYSFSDWYSPAC 127
Db 71 SGGANSPALYLLDGLRAQDDFSQWDINTPAFVWYQSGLSIVMPVGGSSPYSFSDWYOPAC 130
QY 128 GKAGCQTYKWETFLTSELTPQLWLSANRAVKPTGSAAGLSMAGSSAMILAAHYHPQQFIYAG 187
Db 131 GKAGCQTYKWETFLTSELTPQLWLSANRHVKPTGSAVVGLSMAASALTALIVHPQFFYAG 190
QY 188 SLSALLDPSQMGPSLIGLAMGAGGKAAADMGPPSDPAWERNDPQQIPKL VANNTRL 247
Db 191 AMSGLLDPSQAMGFTLIGLAMGAGGKASDMWPKEDPAWQRNDPILNKGKLIANNTRV 250
QY 248 WVYCGNGTPELGCANIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNPFPNGTHSWEYW 307
Db 251 WVYCGNGKPSDLGNLIPAKFLEGFVRTSNIKFQDAYNAGGGHNGVDFPDGTHSWEYW 310
QY 308 GAQLNAMKGDLOSSLGA 324
Db 311 GAQLNAMKPDQLQALGA 327

Query Match      80.6%; Score 1407; DB 11; Length 338;
Best Local Similarity 78.9%; Pred. No. 6.9e-125;
Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

QY 8 IRANGRLMIGTAAAVLPGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGRDQKVFQ 67
Db 11 VTGMSRLVVGAAVLVGLVAVGVTATAGAFSRPGLPVEYLQVPSMGRDQKVFQ 70
QY 68 SGGNNSPAVYLLDGLRAQDDYNGWDINTPAFVWYQSGLSIVMPVGGSSPYSFSDWYSPAC 127
Db 71 SGGANSPALYLLDGLRAQDDFSQWDINTPAFVWYQSGLSIVMPVGGSSPYSFSDWYOPAC 130
QY 128 GKAGCQTYKWETFLTSELTPQLWLSANRAVKPTGSAAGLSMAGSSAMILAAHYHPQQFIYAG 187
Db 131 GKAGCQTYKWETFLTSELTPQLWLSANRHVKPTGSAVVGLSMAASALTALIVHPQFFYAG 190
QY 188 SLSALLDPSQMGPSLIGLAMGAGGKAAADMGPPSDPAWERNDPQQIPKL VANNTRL 247
Db 191 AMSGLLDPSQAMGFTLIGLAMGAGGKASDMWPKEDPAWQRNDPILNKGKLIANNTRV 250
QY 248 WVYCGNGTPELGCANIPAEFLENFVRSSNLKFQDAYNAAGGHNAVFNPFPNGTHSWEYW 307
Db 251 WVYCGNGKPSDLGNLIPAKFLEGFVRTSNIKFQDAYNAGGGHNGVDFPDGTHSWEYW 310
QY 308 GAQLNAMKGDLOSSLGA 324
Db 311 GAQLNAMKPDQLQALGA 327

RESULT 25
US-10-051-643-32
; Sequence 32, Application US/10051643
```



```
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium Vaccae
; FILE REFERENCE: 11000.10082
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 32
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-051-643-32

Query Match      80.6%; Score 1407; DB 14; Length 338;
Best Local Similarity 78.9%; Pred. No. 6.9e-125;
Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

QY 8 IRAGRLMIGTAAAVLVGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGRDIKVFQ 67
Db 11 VTGMSRLVVGAAVLSGLVAVGVTATAGAFSRPGLPVEYLQVPSMGRDIKVFQ 70
QY 68 SGNNSPAVLLDGLRAQDNDYNDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPAC 127
Db 71 SGANSFALYLLDGLRAQDNDYNDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPAC 130
QY 128 GKAGCQTYKWEFTLSELFWLQWLSANRAVKPTGSAALGLSMAGSSAMTLAAHYHPQQFIYAG 187
Db 131 GKAGCQTYKWEFTLSELFWLQWLSANRAVKPTGSAALGLSMAGSSAMTLAAHYHPQQFIYAG 190
QY 188 SLSALLDPSQGMGPTLIGLWAGDAGGYKAADMWGPSSDPAWERNDDPTQIIPKLVANNTRL 247
Db 191 AMSGLLDPSQGMGPTLIGLWAGDAGGYKASDMWGPSSDPAWERNDDPTQIIPKLVANNTRL 250
QY 248 WYCGNGTNEIGGANIPAEFLNFVRSSNLKFDQDAYNAGGHNAPFPNGTHSWEY 307
Db 251 WYCGNGKPSDLGGNNLPKAFLEGFVRTSNIRKFDQDAYNAGGHNAPFPNGTHSWEY 310
QY 308 GAQLNAMKGDLOSSIGA 324
Db 311 GAQLNAMKPDQLRALGA 327

RESULT 26
US-10-051-643-34
; Sequence 34, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium Vaccae
; FILE REFERENCE: 11000.10082
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 34
; LENGTH: 338
; TYPE: PRT
```

```
; ORGANISM: Mycobacterium bovis
US-10-051-643-34

Query Match      80.6%; Score 1407; DB 14; Length 338;
Best Local Similarity 78.9%; Pred. No. 6.9e-125;
Matches 250; Conservative 29; Mismatches 38; Indels 0; Gaps 0;

QY 8 IRAGRLMIGTAAAVLVGLVGLAGGAATAGAFSRPGLPVEYLQVPSMGRDIKVFQ 67
Db 11 VTGMSRLVVGAAVLSGLVAVGVTATAGAFSRPGLPVEYLQVPSMGRDIKVFQ 70
QY 68 SGNNSPAVLLDGLRAQDNDYNDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPAC 127
Db 71 SGANSFALYLLDGLRAQDNDYNDINTPAFEWYQSGLSIVMPVGGQSSFYSDWYSPAC 130
QY 128 GKAGCQTYKWEFTLSELFWLQWLSANRAVKPTGSAALGLSMAGSSAMTLAAHYHPQQFIYAG 187
Db 131 GKAGCQTYKWEFTLSELFWLQWLSANRAVKPTGSAALGLSMAGSSAMTLAAHYHPQQFIYAG 190
QY 188 SLSALLDPSQGMGPTLIGLWAGDAGGYKAADMWGPSSDPAWERNDDPTQIIPKLVANNTRL 247
Db 191 AMSGLLDPSQGMGPTLIGLWAGDAGGYKASDMWGPSSDPAWERNDDPTQIIPKLVANNTRL 250
QY 248 WYCGNGTNEIGGANIPAEFLNFVRSSNLKFDQDAYNAGGHNAPFPNGTHSWEY 307
Db 251 WYCGNGKPSDLGGNNLPKAFLEGFVRTSNIRKFDQDAYNAGGHNAPFPNGTHSWEY 310
QY 308 GAQLNAMKGDLOSSIGA 324
Db 311 GAQLNAMKPDQLRALGA 327

RESULT 27
US-09-051-755-17
; Sequence 17, Application US/09051755
; Patent No. US20010018048A1
; GENERAL INFORMATION:
; APPLICANT: LEER, Robert J
; APPLICANT: POWELS, Patricia L
; APPLICANT: CONWAY, Patricia L
; TITLE OF INVENTION: NOVEL ADHERENCE FACTORS OF NON PATHOGENIC
; TITLE OF INVENTION: MICROORGANISMS AND APPLICATIONS THEREOF FOR SCREENING
; TITLE OF INVENTION: MICROORGANISMS FOR SPECIFIC PROBIOTIC PROPERTIES; NOVEL
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS AND FOOD ADDITIVES C
; FILE REFERENCE: LEER et al. 09/051,755
; CURRENT APPLICATION NUMBER: US/09/051,755
; CURRENT FILING DATE: 1998-09-30
; EARLIER APPLICATION NUMBER: PCT/NL95/00367
; EARLIER FILING DATE: 1995-10-20
; EARLIER APPLICATION NUMBER: PCT/NL96/00409
; EARLIER FILING DATE: 1996-10-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:aligned aa
; OTHER INFORMATION: sequence of Mycobacterium Mlep85b
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (1)..(364)
; OTHER INFORMATION: Xaa is any amino acid.
US-09-051-755-17

Query Match      76.8%; Score 1339.5; DB 9; Length 364;
Best Local Similarity 77.2%; Pred. No. 1.9e-118;
Matches 268; Conservative 19; Mismatches 35; Indels 25; Gaps 11;

QY 1 MTDVSRKIRA-----WGRIMIGTAAAVLVGLVGLAGGAATAGAFSRPG-LPVEYLQVPS 55
Db 3 MIDVSGKIRAXXXXKWRWLLVGAATAATXXLPSLSIAGGAATAAFSRPGKLPVEYLQVPS 62
```

Qy 56 P-SWGRDIKV-QFQSGGNNSPAVYLLDGLRAQDDY--NGWDINTPAPEWYQSG-LSIVM 110
Db 63 EXAMGRTIKVQFQNGGNGSPAVYLLDGLRAQDDYXXNGWDINTSAPEWYQSGXLSVVM 122
Qy 111 PVGQSSFYSDWSPACGKAGCTTYKWTETFLTSELPMWLSANRAVKTGSAALGL---- 165
Db 123 PVGQSSFYSDWSPACGKAGCTTYKWTETFLTSELPMWLSANRAVKTGSAALGLXXXX 182
Qy 166 --SWAGSS---AMILAAHPQOFI---YAGSLALLDPSQGMGSPSLIGLAMGDAGGYKAA 217
Db 183 XXSWAGSSXXXALILAAHPDQFIXXYAGSLALLDPSQGMGSPSLIGLAMGDAGGYKAA 242
Qy 218 DMWGSPSDPAWERNPTQOIPKLVANTRLWVYCGNTPNELGANTPAEFLENFVRSN 277
Db 243 DMWGPPNDPAWQRNDPILQAGKLVANNTLWVYCGNTPSELGNTNPAEFLENFVHGSN 302
Qy 278 LKQDAYNAAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 322
Db 303 LKQDAYNAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 349
Qy 278 LKQDAYNAAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 322
Db 303 LKQDAYNAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 349

RESULT 28

US-10-231-055-17
; Sequence 17, Application US/10231055
; Publication No. US20030095978A1
; GENERAL INFORMATION:
; APPLICANT: LEER, Robert J
; APPLICANT: POWELS, Patricia L
; TITLE OF INVENTION: NOVEL ADHERENCE FACTORS OF NON PATHOGENIC
; TITLE OF INVENTION: MICROORGANISMS AND APPLICATIONS THEREOF FOR SCREENING
; TITLE OF INVENTION: MICROORGANISMS FOR SPECIFIC PROBIOTIC PROPERTIES; NOVEL
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS AND FOOD ADDITIVES C
; FILE REFERENCE: LEER et al. 09/051,755
; CURRENT FILING DATE: 2002-08-30
; PRIOR FILING DATE: 2002-08-30
; PRIOR FILING DATE: 1998-09-30
; PRIOR FILING DATE: 1995-10-20
; PRIOR FILING DATE: 1996-10-21
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 17
; LENGTH: 364
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: aligned aa
; OTHER INFORMATION: sequence of Mycobacterium Mlep85b

Qy 1 MTDVSRKTRA---WGRRLMTGTAADVPLGLVGLAGGAATAGAFSPG-LPVEYLOVPS 55
Db 3 MIDVSGKIRAXXXKRWGLLVGAATXXLPGLSLIAGGAATAGAFSPG-LPVEYLOVPS 62
Qy 56 P-SWGRDIKV-QFQSGGNNSPAVYLLDGLRAQDDY--NGWDINTPAPEWYQSG-LSIVM 110
Db 63 EXAMGRTIKVQFQNGGNGSPAVYLLDGLRAQDDYXXNGWDINTSAPEWYQSGXLSVVM 122
Qy 111 PVGQSSFYSDWSPACGKAGCTTYKWTETFLTSELPMWLSANRAVKTGSAALGL---- 165
Db 123 PVGQSSFYSDWSPACGKAGCTTYKWTETFLTSELPMWLSANRAVKTGSAALGLXXXX 182
Qy 166 --SWAGSS---AMILAAHPQOFI---YAGSLALLDPSQGMGSPSLIGLAMGDAGGYKAA 217
Db 183 XXSWAGSSXXXALILAAHPDQFIXXYAGSLALLDPSQGMGSPSLIGLAMGDAGGYKAA 242
Qy 218 DMWGSPSDPAWERNPTQOIPKLVANTRLWVYCGNTPNELGANTPAEFLENFVRSN 277
Db 243 DMWGPPNDPAWQRNDPILQAGKLVANNTLWVYCGNTPSELGNTNPAEFLENFVHGSN 302
Qy 278 LKQDAYNAAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 322
Db 303 LKQDAYNAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 349

Query Match

Best Local Similarity 76.8%; Score 1339.5; DB 15; Length 364;
Matches 268; Conservative 19; Mismatches 35; Indels 25; Gaps 11;

Qy 1 MTDVSRKTRA---WGRRLMTGTAADVPLGLVGLAGGAATAGAFSPG-LPVEYLOVPS 55
Db 3 MIDVSGKIRAXXXKRWGLLVGAATXXLPGLSLIAGGAATAGAFSPG-LPVEYLOVPS 62
Qy 56 P-SWGRDIKV-QFQSGGNNSPAVYLLDGLRAQDDY--NGWDINTPAPEWYQSG-LSIVM 110
Db 63 EXAMGRTIKVQFQNGGNGSPAVYLLDGLRAQDDYXXNGWDINTSAPEWYQSGXLSVVM 122
Qy 111 PVGQSSFYSDWSPACGKAGCTTYKWTETFLTSELPMWLSANRAVKTGSAALGL---- 165
Db 123 PVGQSSFYSDWSPACGKAGCTTYKWTETFLTSELPMWLSANRAVKTGSAALGLXXXX 182

Qy 166 --SWAGSS---AMILAAHPQOFI---YAGSLALLDPSQGMGSPSLIGLAMGDAGGYKAA 217
Db 183 XXSWAGSSXXXALILAAHPDQFIXXYAGSLALLDPSQGMGSPSLIGLAMGDAGGYKAA 242
Qy 218 DMWGSPSDPAWERNPTQOIPKLVANTRLWVYCGNTPNELGANTPAEFLENFVRSN 277
Db 243 DMWGPPNDPAWQRNDPILQAGKLVANNTLWVYCGNTPSELGNTNPAEFLENFVHGSN 302
Qy 278 LKQDAYNAAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 322
Db 303 LKQDAYNAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 349
Qy 278 LKQDAYNAAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 322
Db 303 LKQDAYNAGGHNVAFFPNPNGTHSWE--YWGAQLNAMKGDQLQSSL 349

RESULT 29

US-09-880-505-30
; Sequence 30, Application US/09880505
; Publication No. US20030007976A1
; GENERAL INFORMATION:
; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L.J.
; APPLICANT: Prestidge, Ross
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: Of Immunologically-Mediated Skin Disorders
; FILE REFERENCE: 11000.1007C2
; CURRENT APPLICATION NUMBER: US/09/880,505
; CURRENT FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: US 09/324,542
; PRIOR FILING DATE: 1999-06-02
; PRIOR APPLICATION NUMBER: US 08/997,080
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-09-880-505-30

Query Match 76.5%; Score 1335.5; DB 11; Length 330;
Best Local Similarity 76.6%; Pred. No. 4e-118;
Matches 239; Conservative 31; Mismatches 41; Indels 1; Gaps 1;

Qy 13 RRLMIGTAAAVLPLGLVGLAGGAATAGAFSPG-LPVEYLOVPSMGRIKVKQFQSGNN 72
Db 16 RRLVVEAMGVALLSALIGVY-GSAPAFSPG-LPVEYLOVPSMGRIKVKQFQSGNN 74
Qy 73 SPAYLLDGLRAQDDYNGWDINTPAPEWYQSGLSIYMPVGGQSSFYSDWYSPACGKAGC 132
Db 75 SPALYLLDGLRAQDDYNGWDINTPAPEWYQSGLSIYMPVGGQSSFYSDWYSPACGKAGC 134
Qy 133 QTYKWTETFLTSELPMWLSANRAVKTGSAALIGLSMAGSSAMILAAHPQFIYAGSLAL 192
Db 135 QTYKWTETFLTSELPMWLSANRAVKTGSAALIGLSMAGSSAMILAAHPQFIYAGSLAL 194
Qy 193 LDPSQGMGSPSLIGLAMGDAGGYKAADWMPGSSDDPAWERNPTQOIPKLVANTRLWVYCG 252
Db 195 LDPSNMGSPSLIGLAMGDAGGYKAADWMPGSSDDPAWERNPTQOIPKLVANTRLWVYCG 254
Qy 253 NGTNELEGGANIPAEFLFENFVRSNLFQDAYNAAGGHNVAFFPNPNGTHSWEYWGALN 312
Db 255 NGKTELEGGNLPAKLEGLVRTSNIKFQDGYNAGGHNVAFFPNPNGTHSWEYWGALN 314
Qy 313 AMKGDQLQSSILGA 324
Db 315 DMKPDLOQYLGA 326

RESULT 30

US-10-051-643-30
; Sequence 30, Application US/10051643
; Publication No. US20020197265A1
; GENERAL INFORMATION:

```

; APPLICANT: Watson, James D.
; APPLICANT: Tan, Paul L. J.
; TITLE OF INVENTION: Methods and Compounds for the Treatment
; TITLE OF INVENTION: of Immunologically-Mediated Diseases of the Respiratory
; TITLE OF INVENTION: System using Mycobacterium Vaccae
; FILE REFERENCE: 11000.1008c2
; CURRENT APPLICATION NUMBER: US/10/051,643
; CURRENT FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US09/156,181
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: US 08/996,624
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 208
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 30
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Mycobacterium leprae
US-10-051-643-30

Query Match          76.5%; Score 1335.5; DB 14; Length 330;
Best Local Similarity 76.6%; Pred. No. 4e-118;
Matches 239; Conservative 31; Mismatches 41; Indels 1; Gaps 1;

QY 13 RRLMIGTAAAVLPLGLVGLAGGAATAGAFSPGLPVEYLQVPSMGRDIKVFQSGGNN 72
Db 16 RRLVVEAMGVALLSALIGVV-GSAPAEAFSPGLPVEYLQVPSMGRDIKVFQSGGNN 74

QY 73 SPAYLLDGLRAQDDYNGWDINTPAFEWYQSGLSIVMPVGGSSFYSDWYSPACGKAGC 132
Db 75 SPALYLLDGLRAQDDFGWDINTTAFEWYQSGLSIVMPVGGSSFYSDWYSPACGKAGC 134

QY 133 QTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHYHPQQFIYAGSL 192
Db 135 QTYKWETFLTSELPQWLSANRAVKPTGSAALGLSMAGSSAMILAAHYHPQQFIYAGSL 194

QY 193 LDPSQMGPSLIGLAMGDAGYKAADWMPSSDPAWRNDPTQOI PKLVANNTELWYCG 252
Db 195 LDPSNMGPSLIGLAMGDAGYKAADWMPSSDPAWRNDPTQOI PKLVANNTELWYCG 254

QY 253 NGTPNELGGANI PAEFL ENFVRSNLKFDAYNAAGGHNAVFNP PPNGTHSWEYWGQALN 312
Db 255 NGKPTLGGNNLPKLLLEGLVTSNIRKFDQYNAGGCHNAVFNP PPDSTHSGTHSWEYWGQALN 314

QY 313 AMKGDLSQSLGA 324
Db 315 DMKPDLSQSLGA 326

```

Search completed: February 5, 2004, 17:43:19
Job time: 45.881 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:25:12 ; Search time 16.9643 Seconds
(without alignments)
888.869 Million cell updates/sec

Title: US-09-805-427A-1
Perfect score: 479
Sequence: 1 MTEQWNFAGIEAASAIQ.....ISEAGQMASTEGNVTGMFA 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1078588

Minimum DB seq length: 6
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

Database :			
A_Geneseq_19Jun03:*			
1:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*	100.0	19
2:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*	100.0	20
3:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*	54.1	51
4:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*	54.1	51
5:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*	54.1	51
6:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*	54.1	51
7:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*	54.1	51
8:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*	54.1	51
9:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*	54.1	51
10:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*	54.1	51
11:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*	54.1	51
12:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*	54.1	51
13:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*	54.1	51
14:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*	54.1	51
15:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*	54.1	51
16:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*	54.1	51
17:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*	54.1	51
18:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*	54.1	51
19:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*	54.1	51
20:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*	54.1	51
21:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*	54.1	51
22:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*	54.1	51
23:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*	54.1	51
24:	/SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*	54.1	51

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	100.0	95	16	AAW11494
2	479	100.0	95	20	AAV29890
3	479	100.0	95	20	AAV29788
4	479	100.0	95	22	AB35219
5	479	100.0	95	23	AB30968
6	479	100.0	95	23	AAW50740
7	479	100.0	95	23	AAW50740
8	479	100.0	403	19	AAW72943
9	479	100.0	403	20	AAW21963

10	479	100.0	404	19	AAW72942	Mycobacterium tube
11	479	100.0	404	20	AAV21962	Amino acid sequenc
12	259	54.1	51	18	AAW32466	Mycobacterium tube
13	259	54.1	51	18	AAW32339	Mycobacterium tube
14	259	54.1	51	19	AAW61701	M. tuberculosis im
15	259	54.1	51	19	AAW64334	Mycobacterium tube
16	259	54.1	51	20	AAW39313	Mycobacterium tube
17	259	54.1	51	20	AAV38988	M. tuberculosis ES
18	259	54.1	51	22	AAU01897	M. tuberculosis re
19	259	54.1	51	23	AAE29716	Mycobacterium tube
20	259	54.1	51	23	AAE17580	Mycobacterium tube
21	186	38.8	95	23	ABU05360	Mycobacterium spec
22	151	31.5	96	22	AAW30391	M. tuberculosis an
23	106	22.1	20	18	AAW35547	C glutamicum prote
24	106	22.1	20	18	AAW35548	ESAT-6 antigen SEQ
25	101	21.1	20	21	AAW88580	Antigenic C-termin
26	98	20.5	20	22	AAE12275	Mycobacterium tube
27	98	20.5	20	22	AAE12275	Mycobacterium tube
28	97	20.3	20	22	AAV21964	Amino acid sequenc
29	96	20.0	19	18	AAW35548	Mycobacterium tube
30	87	18.2	17	21	AAW88581	ESAT-6 antigen SEQ
31	86.5	18.1	100	22	AAW35229	Antigenic N-termin
32	85	17.7	96	19	AAW2884	M tuberculosis Rv3
33	85	17.7	96	20	AAV21901	Mycobacterium tube
34	85	17.7	96	22	AAW35220	Amino acid sequenc
35	85	17.7	96	23	ABU05727	M tuberculosis Rv0
36	81.5	17.0	108	22	AAW30390	M. tuberculosis an
37	81	16.9	15	21	AAW94597	C glutamicum prote
38	81	16.9	15	23	ABG30959	Mycobacterium tube
39	81	16.9	15	23	ABG30960	Mycobacterium tube
40	81	16.9	15	23	ABG30961	Mycobacterium tube
41	81	16.9	15	24	ABG75863	Mycobacterium tube
42	81	16.9	96	23	ABU05648	M. tuberculosis an
43	80.5	16.8	90	22	AAW35233	M tuberculosis Rv3
44	80	16.7	317	8	AAW70303	Sequence of flagel
45	80	16.7	369	8	AAW70309	Sequence of flagel
46	80	16.7	387	8	AAW70307	Sequence of flagel
47	80	16.7	395	8	AAW70308	Sequence of flagel
48	80	16.7	414	8	AAW70313	Sequence of flagel
49	80	16.7	425	8	AAW70315	Sequence of flagel
50	80	16.7	447	8	AAW70314	Sequence of flagel
51	80	16.7	498	8	AAW70302	Sequence of flagel
52	80	16.7	498	22	AAW94641	Escherichia coli H
53	80	16.7	500	17	AAW90303	Thioredoxin-flagel
54	79	16.5	15	21	AAW94590	Mycobacterium tube
55	79	16.5	15	21	AAW94592	Mycobacterium tube
56	79	16.5	15	23	ABG30951	Mycobacterium tube
57	79	16.5	15	23	ABG30962	Mycobacterium tube
58	79	16.5	15	23	AAO17432	Early secreted ant
59	78.5	16.4	368	20	AAW32069	Mycobacterium tube
60	78.5	16.4	368	23	AAU74598	Antigenic fusion p
61	78.5	16.4	600	23	AAW32068	Mycobacterium tube
62	78.5	16.4	600	23	AAU74597	Antigenic fusion p
63	78	16.3	15	21	AAW94596	Mycobacterium tube
64	78	16.3	15	23	ABG30957	M tuberculosis Rv3
65	77.5	16.2	125	22	AAW35230	Mycobacterium tube
66	77	16.1	15	21	AAW94595	Mycobacterium tube
67	77	16.1	15	23	ABG30952	Mycobacterium tube
68	77	16.1	15	23	AAO17433	Early secreted ant
69	77	16.1	876	23	ABP65529	Bifidobacterium lo
70	76.5	16.0	837	22	AAU34387	Staphylococcus aur
71	76.5	16.0	875	22	AAU34387	Staphylococcus aur
72	76.5	16.0	2434	22	AAU34339	Staphylococcus aur
73	76.5	16.0	6281	22	AAU34703	Staphylococcus aur
74	76.5	16.0	10498	24	ABU19119	Staphylococcus aur
75	76	15.9	15	23	ABG30958	Pathogen specific
76	76	15.9	16	22	AAE12278	Mycobacterium tube
77	76	15.9	424	22	AAU47141	Mycobacterium tube
78	74	15.4	15	21	AAW94598	Propionibacterium
79	74	15.4	15	23	ABG30955	Mycobacterium tube
80	74	15.4	15	23	ABG30967	Mycobacterium tube
81	74	15.4	166	22	ABG04606	Novel human diagno
82	73.5	15.3	772	23	ABP30556	Streptococcus poly

83	73.5	15.3	811	23	ABP27026	Streptococcus poly
84	73	15.2	15	23	ABG30956	Mycobacterium tube
85	73	15.2	2086	22	AAU34143	Staphylococcus aur
86	73	15.2	2470	22	ABB61247	Drosophila melanog
87	73	15.2	5795	22	AAU37017	Staphylococcus aur
88	72.5	15.1	299	20	AAV32067	Mycobacterium tube
89	72.5	15.1	299	23	AAE29723	Mycobacterium sp.
90	72.5	15.1	299	23	AAE17587	Mycobacterium spec

RESULT 1

AAW11494

AAW11494 standard; Protein; 95 AA.

AC

AAW11494;

XX

27-MAR-1997 (first entry)

DT

XX

ESAT6.

DE

XX

XX

XX

ESAT6; Mycobacterium tuberculosis; 6kDa antigen; tuberculosis complex;

KW

XX

HYB76-8 reactive antigen; interferon-gamma; memory T-lymphocyte; vaccine;

KW

XX

M. africanum; M. bovis; delayed-type hypersensitivity reaction; therapy.

KW

XX

Mycobacterium tuberculosis.

OS

XX

WO9501441-A1.

PN

XX

12-JAN-1995.

PD

XX

01-JUL-1994; 94WO-DK00273.

PF

XX

02-JUL-1993; 93DK-0000798.

PR

XX

(STAT-) STATENS SERUMINSTITUT.

PA

XX

Andersen AB, Andersen P, Haslov K, Sorensen A;

PI

XX

WPI; 1995-061005/08.

DR

XX

N-PSDB; AAT51422.

DR

XX

Vaccine for tuberculosis induces interferon-gamma release from

PT

XX

T-lymphocytes - comprises an antigen released from mycobacteria,

PT

XX

for immunisation of humans

PT

XX

Claim 5; Page 61-63; 101pp; English.

PS

XX

This sequence represents the Mycobacterium tuberculosis ESAT6. ESAT6 is

CC

XX

also known as the 6kDa antigen, or the HYB76-8 reactive antigen. ESAT6

CC

XX

is released from metabolising bacteria, and can be isolated from short

CC

XX

term filtrates grown as shaken cultures for 7 days. ESAT6 also induces a

CC

XX

release of interferon-gamma from reactivated memory T-lymphocytes. This

CC

XX

protein sequence is included in the vaccine of the invention. The

CC

XX

vaccine is for immunising an animal (including humans) against

CC

XX

tuberculosis caused by a Mycobacterium belonging to the tuberculosis

CC

XX

complex. The Mycobacterium of the tuberculosis complex are, M.

CC

XX

tuberculosis, M. africanum, and M. bovis. The vaccine can evoke a

CC

XX

protective immune response against tuberculosis or a delayed-type

CC

XX

hypersensitivity reaction. The protein can also be included in a

CC

XX

composition for diagnosing tuberculosis. The composition is injected

CC

XX

intradermally, and a skin reaction is an indicator of tuberculosis.

CC

XX

Sequence 95 AA;

XX

XX

Query Match 100.0%; Score 479; DB 16; Length 95;

Best Local Similarity 100.0%; Pred. No. 3.3e-44;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1

MTEQWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGSGSEAYQGVQKWD

60

Db	1	MTEQWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGSGSEAYQGVQKWD	60
Qy	61	TATELNALONLARTISEAGQAMASTEGNVTGMEFA	95
Db	61	TATELNALONLARTISEAGQAMASTEGNVTGMEFA	95

RESULT 2

AAV29890

AAV29890 standard; Protein; 95 AA.

ID

XX

AAV29890;

AC

XX

18-NOV-1999 (first entry)

DT

XX

Mycobacterium tuberculosis ESAT-6 protein sequence.

DE

XX

XX

XX

XX

Mycobacterial; lactic acid bacterium; diagnosis; skin test; vaccine;

KW

XX

delayed type hypersensitivity; DTH; ESAT-6 homodimer; tuberculosis;

KW

XX

interferon-gamma release.

KW

XX

Mycobacterium tuberculosis.

OS

XX

WO9945119-A2.

PN

XX

10-SEP-1999.

PD

XX

05-MAR-1999; 99WO-DK00109.

PF

XX

06-MAR-1998; 98DK-0000306.

PR

XX

06-MAR-1998; 98US-0077105.

PR

XX

(STAT-) STATENS SERUM INST.

PA

XX

Jensen CL, Folkersen J;

PI

XX

WPI; 1999-551043/46.

DR

XX

N-PSDB; AAZ21132.

DR

XX

New mycobacterial polypeptide produced in lactic acid bacteria, useful

PT

XX

in tuberculosis diagnosis and vaccines -

PT

XX

Disclosure; Page 75-76; 76pp; English.

PS

XX

The present invention describes a bioactive polypeptide (or

CC

XX

immunologically equivalent analogue) produced in lactic acid bacteria

CC

XX

which reacts with lymphoid cells primed with Mycobacterium tuberculosis

CC

XX

complex mycobacteria (M. tuberculosis, M. africanum or M. bovis). The

CC

XX

polypeptide and ESAT-6 polypeptides are useful in compositions for

CC

XX

diagnosis of and vaccination against tuberculosis caused by

CC

XX

M. tuberculosis complex mycobacteria. The ESAT-6 polypeptide can be used

CC

XX

to diagnose ongoing/previous sensitisation with these bacteria by

CC

XX

detecting cytokine release when contacting blood samples with the

CC

XX

polypeptide. The bioactive polypeptide may be used in diagnostic

CC

XX

compositions and vaccines for mycobacteria other than of the

CC

XX

M. tuberculosis complex, e.g. M. avium which infects poultry and

CC

XX

occasionally humans M. leprae; they are especially useful when they do

CC

XX

not react with lymphoid cells previously primed with M. tuberculosis

CC

XX

complex mycobacteria, and so do not give rise to a diagnostic reaction

CC

XX

in individuals infected with these bacteria. The polypeptides may also

CC

XX

be used in in vitro diagnostic tests e.g. stimulation of interferon-gamma

CC

XX

release from lymphocytes. The polypeptide has similar or higher

CC

XX

bioactivity as currently used tuberculin reagent in the standard

CC

XX

delayed type hypersensitivity (DTH) skin test for tuberculosis, but may

CC

XX

have greater specificity, being better able to discriminate between

CC

XX

lymphoid cells primed from tuberculosis and from previous vaccination.

CC

XX

The present sequence represents M. tuberculosis ESAT-6 used in the

CC

XX

exemplification of the present invention.

CC

XX

Sequence 95 AA;

XX

XX

Query Match 100.0%; Score 479; DB 20; Length 95;

Best Local Similarity 100.0%; Pred. No. 3.3e-44;

Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAYQGVQKQWDA 60
 Db 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAYQGVQKQWDA 60
 QY 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95
 Db 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95

RESULT 3
 AAY29788
 ID AAY29788 standard; Protein; 95 AA.
 XX
 AC AAY29788;
 XX
 DT 08-NOV-1999 (first entry)
 XX
 DE Mycobacterial tuberculosis ESAT-6 protein.
 XX
 KW Mycobacterial; ESAT-6; 6 kDa antigen; tuberculosis; infection;
 KW immunological response; diagnosis; vaccine.
 XX
 OS Mycobacterium tuberculosis.
 XX
 FN US5955077-A.
 XX
 PD 21-SEP-1999.
 XX
 PF 05-JUN-1995; 95US-0465640.
 XX
 PR 05-JUN-1995; 95US-0465640.
 PR 20-SEP-1993; 93US-0123182.
 PR 01-JUL-1994; 94WO-DK00273.
 XX
 PA (STAT-) STATENS SERUMINSTITUT.
 XX
 PI Andersen AB, Andersen P, Haslov K, Sorensen AL;
 XX
 DR WPI; 1999-539545/45.
 DR N-PSDB; AAZ08877.
 XX
 PT Polypeptide secreted from Mycobacterium is useful as a vaccine
 PT against tuberculosis
 XX
 PS Claim 24; Fig 10C; 39pp; English.
 XX
 CC The present invention describes a purified or non-naturally occurring
 CC polypeptide (I) released from a metabolising mycobacteria comprising an
 CC ESAT6, also called the 6 kDa antigen. The present sequence represents
 CC Mycobacterium tuberculosis ESAT-6 protein. Also described is a purified
 CC or non-naturally occurring polypeptide (II) with the ability to elicit
 CC a delayed type hypersensitivity reaction which comprises a T cell
 CC epitope of (I). (II) can be used with a carrier or vehicle in a
 CC composition for diagnosing tuberculosis caused by mycobacteria belonging
 CC to the tuberculosis complex, i.e. Mycobacterium tuberculosis, M. bovis
 CC and M. africanum. The composition can be used to detect microbial
 CC antibodies or components of mycobacteria in samples or in animals
 CC through the use of immunoassays. (II) can be used as a vaccine for
 CC immunising an animal, including humans against tuberculosis caused by
 CC mycobacteria of the tuberculosis-complex. (II) induce a release of
 CC IFN-gamma from reactivated T-lymphocytes evoking a protective immune
 CC response. Vaccine containing (II) has the same protective potency as
 CC the live BCG vaccine against tuberculosis.
 XX
 SQ Sequence 95 AA;

Query Match 100.0%; Score 479; DB 20; Length 95;
 Best Local Similarity 100.0%; Pred. No. 3.3e-44;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAYQGVQKQWDA 60
 Db 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAYQGVQKQWDA 60
 QY 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95
 Db 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95

RESULT 5
 ABG30968
 ID ABG30968 standard; Protein; 95 AA.
 XX
 AC ABG30968;
 XX

Db 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAYQGVQKQWDA 60
 QY 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95
 Db 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95

RESULT 4
 AAB35219
 ID AAB35219 standard; Protein; 95 AA.
 XX
 AC AAB35219;
 XX
 DT 24-APR-2001 (first entry)
 XX
 DE M tuberculosis Rv3875 protein.
 XX
 KW Tuberculosis; TB; vaccine; esat-6 gene family; Rv0287; Rv1036c;
 KW Rv1037c; Rv2348c; Rv2348c; Rv2653c; Rv2654c; Rv3020c; Rv3444c;
 KW Rv3445c; Rv3890c; Rv3891c; Rv3904c; Rv3905c.
 XX
 OS Mycobacterium tuberculosis.
 XX
 FN WO200104151-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 13-JUL-2000; 2000WO-DK00398.
 XX
 PR 13-JUL-1999; 99DK-0001020.
 PR 15-JUL-1999; 99US-0144011.
 XX
 PA (STAT-) STATENS SERUM INST.
 XX
 PI Andersen P, Skjot R;
 XX
 DR WPI; 2001-091923/10.
 XX
 PT New polypeptide encoded by a member of the esat-6 gene family for
 PT immunizing against and diagnosis of tuberculosis -
 XX
 PS Example 2; Page 65; 80pp; English.
 XX
 CC The present invention provides the protein and coding sequences for
 CC members of the esat-6 gene family from Mycobacterium tuberculosis. These
 CC proteins include Rv0287, Rv1036c, Rv1037c, Rv2348c, Rv2348c, Rv2653c,
 CC Rv2654c, Rv3020c, Rv3444c, Rv3445c, Rv3890c, Rv3891c, Rv3904c and
 CC Rv3905c. These can be used to produce vaccines against, and in the
 CC diagnosis of, tuberculosis (TB) infection. The present sequence is one of
 CC the proteins of the invention.
 XX
 SQ Sequence 95 AA;

Query Match 100.0%; Score 479; DB 22; Length 95;
 Best Local Similarity 100.0%; Pred. No. 3.3e-44;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAYQGVQKQWDA 60
 Db 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAYQGVQKQWDA 60
 QY 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95
 Db 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95

RESULT 5
 ABG30968
 ID ABG30968 standard; Protein; 95 AA.
 XX
 AC ABG30968;
 XX

DT 21-OCT-2002 (first entry)
 XX Mycobacterium tuberculosis ESAT-6.
 DE
 XX
 XX
 KW ESAT-6; mycobacterial infection; tuberculosis; pulmonary tuberculosis;
 KW tuberculosis lymphadenitis; extrapulmonary tuberculosis.
 XX
 OS Mycobacterium tuberculosis.
 XX WO200254072-A2.
 XX
 XX
 XX 11-JUL-2002.
 XX
 XX 08-JAN-2002; 2002WO-GB00055.
 XX
 XX 08-JAN-2001; 2001GB-0000432.
 XX 08-JAN-2001; 2001US-259868P.
 XX
 XX (ISIS-) ISIS INNOVATION LTD.
 XX
 XX Lalvani A;
 XX
 XX WPI; 2002-583633/62.
 XX
 XX Determining the progress of a mycobacterial infection, by direct ex
 XX vivo quantitation of ESAT-6-specific T cells -
 XX
 XX Disclosure; Page 42; 53pp; English.
 XX
 XX The invention describes a method of determining the efficacy of treatment
 XX for mycobacterial infection (such as pulmonary tuberculosis, tuberculosis
 XX lymphadenitis and extrapulmonary tuberculosis). The method involves
 XX determining the level of T cells specific for a mycobacterial antigen
 XX that has decreased after the treatment and therefore determining the
 XX efficacy of the treatment. The method is useful for determining the
 XX efficacy of treatment for mycobacterial infection, the mycobacterial
 XX infection is Mycobacterium tuberculosis or M. bovis infection. The
 XX invention also describes a method useful for determining the presence of
 XX a latent infection in a sample from the individual for the presence of T
 XX cells specific for a mycobacterial antigen. Also described in a method
 XX for determining the effect of an intervention on a mycobacterial
 XX infection in an individual an a method for treating an individual
 XX infected by a mycobacterium. This is the amino acid sequence of
 XX Mycobacterium tuberculosis ESAT-6, the mycobacterial antigen measured in
 XX the invention.
 XX
 XX Sequence 95 AA;
 XX
 XX Query Match 100.0%; Score 479; DB 23; Length 95;
 XX Best Local Similarity 100.0%; Pred. No. 3.3e-44;
 XX Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 60
 DB 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 60
 QY 61 TATELNALQNLARTISEAGQWASTEGNVTGMFA 95
 DB 61 TATELNALQNLARTISEAGQWASTEGNVTGMFA 95
 RESULT 6
 AAM50740
 ID AAM50740 standard; Protein; 95 AA.
 XX
 XX AAM50740;
 XX
 XX 18-APR-2002 (first entry)
 XX
 XX Mycobacterium tuberculosis immunodominant Mtb protein ESAT6.
 XX Mtb; ESAT6; immunogen; mycobacteria; immunisation; vaccine.
 XX

OS Mycobacterium tuberculosis.
 XX WO200204018-A2.
 XX
 XX 17-JAN-2002.
 XX
 XX 10-JUL-2001; 2001WO-US21717.
 XX
 XX 10-JUL-2000; 2000US-217646P.
 XX
 XX (COLS) UNIV COLORADO STATE RES FOUND.
 XX Orme IM, Belisle JT;
 XX WPI; 2002-164602/21.
 XX
 XX Vaccine for boosting immunity to mycobacteria when administered in
 XX mid-life in a subject who has been vaccinated in childhood with
 XX Bacillus Calmette-Guerrin, has purified proteins from mycobacterium
 XX tuberculosis -
 XX
 XX Claim 8; Page 18; 61pp; English.
 XX
 XX The present sequence is that of the Mycobacterium tuberculosis
 XX (Mtb) strain H37Rv gene Rv3875 product, designated ESAT6. This
 XX is one of 31 immunodominant secreted or cytosolic Mtb proteins
 XX of strain H37Rv (see AAM50729-59) discovered through the use of
 XX 2-dimensional liquid phase electrophoresis coupled with an in vitro
 XX interferon-gamma assay and liquid chromatography-mass spectrometry.
 XX The immunogens stimulate a strong interferon-gamma response from
 XX T cells of M. tuberculosis infected mice. The invention provides
 XX vaccine compositions for boosting immunity to mycobacteria when
 XX administered in mid-life to a subject who has been vaccinated
 XX neonatally or in early childhood with BCG and in whom protective
 XX immunity has waned. The vaccine compositions comprise 1 or more
 XX of the 31 purified immunogenic proteins. When used as immunogens,
 XX the secreted Mtb proteins lack the secreted signal sequence. A
 XX preferred protein is Ag85A (see AAM50759), the secreted product of
 XX the Rv3084v gene.
 XX
 XX Sequence 95 AA;
 XX
 XX Query Match 100.0%; Score 479; DB 23; Length 95;
 XX Best Local Similarity 100.0%; Pred. No. 3.3e-44;
 XX Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 60
 DB 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 60
 QY 61 TATELNALQNLARTISEAGQWASTEGNVTGMFA 95
 DB 61 TATELNALQNLARTISEAGQWASTEGNVTGMFA 95
 RESULT 7
 AAY29888
 ID AAY29888 standard; Protein; 196 AA.
 XX
 XX AAY29888;
 XX
 XX 18-NOV-1999 (first entry)
 XX
 XX Mycobacterium tuberculosis ESAT-6 protein sequence.
 XX
 XX Mycobacterial; lactic acid bacterium; diagnosis; skin test; vaccine;
 XX delayed type hypersensitivity; DTH; ESAT-6 homodimer; tuberculosis;
 XX interferon-gamma release.
 XX
 XX Mycobacterium tuberculosis.
 XX
 XX WO9945119-A2.
 XX

PD 10-SEP-1999.
 XX 05-MAR-1999; 99WO-DK00109.
 XX 06-MAR-1998; 98DK-0000306.
 PR 06-MAR-1998; 98US-0077105.
 XX (STAT-) STATENS SERUM INST.
 XX Jensen CL, Folkersen J;
 XX WPI; 1999-551043/46.
 DR N-ESDB; AAZ21131.
 XX New mycobacterial polypeptide produced in lactic acid bacteria, useful
 in tuberculosis diagnosis and vaccines -
 XX Disclosure; Page 74; 76pp; English.
 XX The present invention describes a bioreactive polypeptide (or
 immunologically equivalent analogue) produced in lactic acid bacteria
 which reacts with lymphoid cells primed with Mycobacterium tuberculosis
 complex mycobacteria (M. tuberculosis, M. africanum or M. bovis). The
 polypeptide and ESAT-6 polypeptides are useful in compositions for
 diagnosis of and vaccination against tuberculosis caused by
 M. tuberculosis complex mycobacteria. The ESAT-6 polypeptide can be used
 to diagnose ongoing/previous sensitisation with these bacteria by
 detecting cytokine release when contacting blood samples with the
 polypeptide. The bioreactive polypeptide may be used in diagnostic
 compositions and vaccines for mycobacteria other than of the
 M. tuberculosis complex, e.g. M. avium which infects poultry and
 occasionally humans, M. leprae; they are especially useful when they do
 not react with lymphoid cells previously primed with M. tuberculosis
 complex mycobacteria, and so do not give rise to a diagnostic reaction
 in individuals infected with these bacteria. The polypeptides may also
 be used in in vitro diagnostic tests e.g. stimulation of interferon-gamma
 release from lymphocytes. The polypeptide has similar or higher
 bioreactivity as currently used tuberculin reagent in the standard
 delayed type hypersensitivity (DTH) skin test for tuberculosis, but may
 have greater specificity, being better able to discriminate between
 lymphoid cells primed from tuberculosis and from previous vaccination.
 CC The present sequence represents M. tuberculosis ESAT-6 used in the
 CC exemplification of the present invention.
 XX
 XX Sequence 196 AA;
 SQ
 Query Match 100.0%; Score 479; DB 20; Length 196;
 Best Local Similarity 100.0%; Pred. No. 8.8e-44;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTEQQWNEAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 60
 Db 5 MTEQQWNEAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 64
 QY 61 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 95
 Db 65 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 99
 RESULT 8
 AA72943
 ID AA72943 standard; Protein; 403 AA.
 XX
 XX AA72943;
 AC
 XX 21-JAN-1999 (first entry)
 DT
 XX Mycobacterium tuberculosis antigen ESAT6-MPT59.
 DE
 XX Mycobacterium tuberculosis; antigen; vaccine; immunological;
 KW Mycobacterium tuberculosis; infection.
 XX
 XX Mycobacterium tuberculosis.
 OS

XX WO9844119-A1.
 PN
 XX 08-OCT-1998.
 PD
 XX 01-APR-1998; 98WO-DK00132.
 PF
 XX 05-JAN-1998; 98US-0070488.
 PR 02-APR-1997; 97DK-0000376.
 PR 18-APR-1997; 97US-0044624.
 PR 10-NOV-1997; 97DK-0001277.
 XX (STAT-) STATENS SERUM INST.
 PA
 XX Andersen P, Florio W, Nielsen R, Oettinger T, Rasmussen PB;
 PI Rosenkrands I, Weldingh K;
 PI WPI; 1998-542705/46.
 DR
 XX New isolated mycobacteria polypeptides and nucleic acids - used for
 developing products for the diagnosis of or vaccination against
 mycobacterial infections, particularly tuberculosis
 XX Disclosure; Page 233-234; 163pp; English.
 XX The present sequence represents a Mycobacterium tuberculosis protein.
 CC Products from the present invention, which describes protein fragments
 and nucleic acid fragments derived from M. tuberculosis, can be used in
 CC the detection of and prevention of mycobacterial infections. In
 CC particular, the proteins and nucleic acids can be used for the diagnosis
 CC of or vaccination against tuberculosis caused by M. tuberculosis,
 CC M. africanum or M. bovis.
 XX
 XX Sequence 403 AA;
 SQ
 Query Match 100.0%; Score 479; DB 19; Length 403;
 Best Local Similarity 100.0%; Pred. No. 2.3e-43;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTEQQWNEAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 60
 Db 22 MTEQQWNEAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 81
 QY 61 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 95
 Db 82 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 116
 RESULT 9
 AA21963
 ID AA21963 standard; Protein; 403 AA.
 XX
 XX AA21963;
 AC
 XX 06-SEP-1999 (first entry)
 DT
 XX Amino acid sequence of antigen ESAT-6.
 DE
 XX Immunogenic; Mycobacterium tuberculosis; immune response; infection;
 KW tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
 KW pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
 KW CFP7B; CFP19; CFP27; CFP30A; RD1-ORF; CFP10A; CFP16; CFP19; CFP23;
 KW CFP25A; CFP30B; CFP7B.
 XX Mycobacterium tuberculosis.
 OS
 XX WO924577-A1.
 PN
 XX 20-MAY-1999.
 PD
 XX 08-OCT-1998; 98WO-DK00438.
 PF
 XX 01-APR-1998; 98WO-DK00132.
 PR

```

PR 10-NOV-1997; 97DK-0001277.
PR 05-JAN-1998; 98US-0070488.
XX
XX
PA (STAT-) STATENS SERUM INST.
XX
XX Andersen P, Skjot R;
XX
XX WPI; 1999-347282/29.
XX
XX New immunogenic fragment of Mycobacterium tuberculosis
XX
XX Examples; Page 250-251; 265pp; English.
XX
CC The invention describes a substantially pure immunogenic polypeptide
CC fragment (I) from Mycobacterium tuberculosis that is able to evoke a
CC protective immune response against infections by mycobacteria belonging
CC to the tuberculosis complex. The invention provides a (1) fusion
CC polypeptide comprising at least one polypeptide fragment (I) and at least
CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
CC epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
CC different amino acid sequence from M. tuberculosis, and/or including a
CC sequence which protects the first amino acid sequence from in vivo
CC degradation or post-translational processing; (3) a nucleic acid fragment
CC that encodes the above polypeptides. The polypeptides and nucleic acid
CC are useful as pharmaceuticals for diagnosis of and as antigens for
CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or
CC bovis. The polypeptides are also useful for diagnosing ongoing or
CC previous sensitization in an animal with bacteria belonging to the
CC tuberculosis complex. The invention also describes the use of CFP7A or
CC CFP30A or a T-cell epitope of for the induction of a strong immune
CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell
CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
CC test; use of CFP27, CFP30A, RDI-ORF2, RDI-ORF5, MPT59-ESAT6,
CC ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
CC cell epitope of for the preparation of an immunological composition; and
CC for the preparation of a subunit vaccine.
XX
SQ Sequence 403 AA;
Query Match 100.0%; Score 479; DB 20; Length 403;
Best Local Similarity 100.0%; Pred. No. 2.3e-43;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKSLTKLAAGGSGSEAYQGVQOKWDA 60
DB 22 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKSLTKLAAGGSGSEAYQGVQOKWDA 81

QY 61 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 95
DB 82 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 116

RESULT 10
AAW72942
ID AAW72942 standard; Protein; 404 AA.
XX
XX AAW72942;
AC
XX
XX 21-JAN-1999 (first entry)
DE
DE Mycobacterium tuberculosis antigen MPT59-ESAT6.
XX
XX Mycobacterium tuberculosis; antigen; vaccine; immunological;
KW
KW immunogen; infection.
XX
XX Mycobacterium tuberculosis.
OS
XX
XX WO9844119-A1.
PN
XX
XX 08-OCT-1998.
FD
XX
XX 01-APR-1998; 98WO-DK00132.
PF
XX
XX
XX

```

```

PR 05-JAN-1998; 98US-0070488.
PR 02-APR-1997; 97DK-0000376.
PR 18-APR-1997; 97US-0044624.
PR 10-NOV-1997; 97DK-0001277.
XX
XX (STAT-) STATENS SERUM INST.
XX
XX Andersen P, Florio W, Nielsen R, Oettinger T, Rasmussen PB;
XX Rosenkrands I, Weidigh K;
XX
XX WPI; 1998-542705/46.
XX
XX New isolated mycobacteria polypeptides and nucleic acids - used for
XX developing products for the diagnosis of or vaccination against
XX mycobacterial infections, particularly tuberculosis
XX
XX Disclosure; Page 232-233; 163pp; English.
XX
CC The present sequence represents a Mycobacterium tuberculosis protein.
CC Products from the present invention, which describes protein fragments
CC and nucleic acid fragments derived from M. tuberculosis, can be used in
CC the detection of and prevention of mycobacterial infections. In
CC particular, the proteins and nucleic acids can be used for the diagnosis
CC of or vaccination against tuberculosis caused by M. tuberculosis,
CC M. africanum or M. bovis.
XX
XX Sequence 404 AA;
SQ
Query Match 100.0%; Score 479; DB 19; Length 404;
Best Local Similarity 100.0%; Pred. No. 2.3e-43;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKSLTKLAAGGSGSEAYQGVQOKWDA 60
DB 310 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKSLTKLAAGGSGSEAYQGVQOKWDA 369

QY 61 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 95
DB 370 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 404

RESULT 11
AAW21962
ID AAW21962 standard; Protein; 404 AA.
XX
XX AAW21962;
AC
XX
XX 06-SEP-1999 (first entry)
DE
DE Amino acid sequence of antigen MPT59.
XX
XX Immunogenic; Mycobacterium tuberculosis; immune response; infection;
KW tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
KW pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
KW CFP7B; CFP19; CFP27; CFP30A; RDI-ORF; CFP10A; CFP16; CFP19; CFP23;
KW CFP25A; CFP30B; CFP7B.
XX
XX Mycobacterium tuberculosis.
OS
XX
XX WO9824577-A1.
PN
XX
XX 20-MAY-1999.
PD
XX
XX 08-CCT-1998; 98WO-DK00438.
PF
XX
XX 01-APR-1998; 98WO-DK00132.
PR
PR 10-NOV-1997; 97DK-0001277.
PR 05-JAN-1998; 98US-0070488.
XX
XX (STAT-) STATENS SERUM INST.
XX
XX Andersen P, Skjot R;
XX
XX
XX

```

DR WPI; 1999-347282/29.
 XX New immunogenic fragment of Mycobacterium tuberculosis
 PT Examples; Page 249-250; 265pp; English.
 XX The invention describes a substantially pure immunogenic polypeptide
 CC fragment (I) from Mycobacterium tuberculosis that is able to evoke a
 CC protective immune response against infections by mycobacteria belonging
 CC to the tuberculosis complex. The invention provides a (1) fusion
 CC polypeptide comprising at least one polypeptide fragment (I) and at least
 CC one fusion partner; (2) a fusion polypeptide fragment comprising a T-cell
 CC epitope from M. tuberculosis protein ESAT-6, or MPT59 and a second
 CC different amino acid sequence from M. tuberculosis, and/or including a
 CC sequence which protects the first amino acid sequence from in vivo
 CC degradation or post-translational processing; (3) a nucleic acid fragment
 CC that encodes the above polypeptides. The polypeptides and nucleic acid
 CC are useful as pharmaceuticals, for diagnosis of and as antigens for
 CC vaccination against TB caused by Mycobacterium tuberculosis, africanum or
 CC bovis. The polypeptides are also useful for diagnosing ongoing or
 CC previous sensitization in an animal with bacteria belonging to the
 CC tuberculosis complex. The invention also describes the use of CFP7A or
 CC CFP30A or a T-cell epitope of for the induction of a strong immune
 CC response in a mammal; use of CFP7B, CFP19 or MPT59-ESAT6 or a T-cell
 CC epitope of for diagnosis of TB in a mammal by performing a DTH type skin
 CC test; use of CFP27, CFP30A, RDI-ORF2, RDI-ORF3, RDI-ORF5, MPT59-ESAT6,
 CC ESAT6-MPT59, CFP10A, CFP16, CFP19, CFP23, CFP25A, CFP30B, CFP7B or a T-
 CC cell epitope of for the preparation of an immunological composition; and
 CC for the preparation of a subunit vaccine.
 XX
 SQ Sequence 404 AA;

Query Match 100.0%; Score 479; DB 20; Length 404;
 Best Local Similarity 100.0%; Pred. No. 2.3e-43;
 Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQWNPAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGSGSEAYQGVQKWDA 60
 Db 310 MTEQWNPAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGSGSEAYQGVQKWDA 369
 QY 61 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 95
 Db 370 TATELNALQNLARTISEAGQAMASTEIGNVTGMFA 404

RESULT 12
 AAW32466
 ID AAW32466 standard; Protein; 51 AA.
 XX AAW32466;
 XX 09-JAN-1998 (first entry)
 XX Mycobacterium tuberculosis antigen ESAT-6.
 DE Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
 XX skin testing; M.tuberculosis.
 XX Mycobacterium tuberculosis.
 OS
 XX WO9709428-A2.
 XX 13-MAR-1997.
 XX 30-AUG-1996; 96WO-US14674.
 XX 12-JUL-1996; 96US-0680574.
 PR 01-SEP-1995; 95US-0523436.
 PR 22-SEP-1995; 95US-0533634.
 PR 22-MAR-1996; 96US-0620874.
 PR 05-JUN-1996; 96US-0659683.
 XX (CORI-) CORIXA CORP.

XX Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;
 PI Twardzik DR, Vedvick TH;
 XX WPI; 1997-192903/17.
 DR N-PSDB; AAT91529.
 XX New immunogenic polypeptide(s) from Mycobacterium tuberculosis - are
 PT useful in vaccines for prevention or treatment of tuberculosis, also
 PT for diagnosis
 XX Disclosure; Page 135; 168pp; English.
 XX A new immunogenic polypeptide has been developed comprising an
 CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
 CC its variant differing only in conservative substitutions and/or
 CC modifications). The present sequence represents a M.tuberculosis
 CC antigen, ESAT-6. The immunogenic protein, and fusion proteins
 CC containing one or more of the proteins or one of the proteins plus
 CC ESAT-6, are useful in vaccines, preferably when formulated with a
 CC non-specific adjuvant, to induce an immune response against
 CC M.tuberculosis (for treatment or prevention).
 XX Sequence 51 AA;

Query Match 54.1%; Score 259; DB 18; Length 51;
 Best Local Similarity 100.0%; Pred. No. 9.3e-21;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQWNPAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGSGSEAY 51
 Db 1 MTEQWNPAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGSGSEAY 51

RESULT 13
 AAW32339
 ID AAW32339 standard; Protein; 51 AA.
 XX AAW32339;
 XX 13-JAN-1998 (first entry)
 XX Mycobacterium tuberculosis antigen ESAT-6.
 DE Antigen; immunogen; vaccine; tuberculosis; non specific adjuvant;
 XX skin testing; M.tuberculosis.
 XX Mycobacterium tuberculosis.
 OS
 XX WO9709429-A2.
 XX 13-MAR-1997.
 XX 30-AUG-1996; 96WO-US14675.
 XX 12-JUL-1996; 96US-0680573.
 PR 01-SEP-1995; 95US-0523435.
 PR 22-SEP-1995; 95US-0532136.
 PR 22-MAR-1996; 96US-0620280.
 PR 05-JUN-1996; 96US-0658800.
 XX (CORI-) CORIXA CORP.

XX Campos-neto A, Dillon DC, Houghton R, Reed SG, Skeiky YAW;
 PI Twardzik DR, Vedvick TH;
 XX WPI; 1997-192904/17.
 DR N-PSDB; AAT91463.
 XX New immunogenic polypeptide(s) from soluble M. tuberculosis antigens
 PT - useful for diagnosis of M. tuberculosis infection
 XX Claim 43; Page 147; 190pp; English.

XX A new immunogenic polypeptide has been developed comprising an
CC immunogenic part of a soluble Mycobacterium tuberculosis antigen (or
CC its variant differing only in conservative substitutions and/or
CC modifications). The present sequence represents a specifically claimed
CC sequence of the ESAT-6 M.tuberculosis antigen. The immunogenic
CC polypeptide can be used to diagnose M.tuberculosis infection by forming
CC complexes with specific antibodies in the sample. Fragments of DNA
CC encoding the immunogenic polypeptide can be used as diagnostic primers
CC or probes and agents that bind to the antigen, especially monoclonal
CC antibodies or equivalent polyclonal antibodies, are also used for
CC diagnosis.
XX
SQ Sequence 51 AA;

Query Match 54.1%; Score 259; DB 18; Length 51;
Best Local Similarity 100.0%; Pred. No. 9.3e-21;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEQOWNFAGIEAAASAIQGNVTSHSLDEGKQSLTKLAAAWGGSGSEAY 51
Db 1 MTEQOWNFAGIEAAASAIQGNVTSHSLDEGKQSLTKLAAAWGGSGSEAY 51

RESULT 14
AAW81701
ID AAW81701 standard; Protein; 51 AA.
XX
AC AAW81701;
XX
DT 27-JAN-1999 (first entry)
XX
DE M. tuberculosis immunogenic polypeptide ESAT-6.
XX Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;
KW vaccine; pharmaceutical; infection; diagnosis.
XX
OS Mycobacterium tuberculosis.
XX
PN WO9816646-A2.
PD
PD 23-APR-1998.
XX
PF 07-OCT-1997; 97WO-US18293.
XX
PR 13-MAR-1997; 97US-0818112.
PR 11-OCT-1996; 96US-0730510.
XX
PA (CORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX
DR WPI; 1998-261042/23.
DR N-PSDB; AAV64501.
XX
PT Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PT and for diagnosis, treatment and prevention of tuberculosis
XX
PS Disclosure; Page 126; 230pp; English.
XX
CC This sequence represents an immunogenic portion of a soluble
CC Mycobacterium tuberculosis (MT) antigen which can be used in a method
CC for inducing protective immunity against tuberculosis (TB). This
CC sequence can be formulated into vaccines and/or pharmaceutical
CC compositions for immunising against M. tuberculosis infection or may
CC be used for the diagnosis of tuberculosis.
XX
SQ Sequence 51 AA;

Query Match 54.1%; Score 259; DB 19; Length 51;
Best Local Similarity 100.0%; Pred. No. 9.3e-21;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEQOWNFAGIEAAASAIQGNVTSHSLDEGKQSLTKLAAAWGGSGSEAY 51
Db 1 MTEQOWNFAGIEAAASAIQGNVTSHSLDEGKQSLTKLAAAWGGSGSEAY 51

RESULT 15
AAW64334
ID AAW64334 standard; Protein; 51 AA.
XX
AC AAW64334;
XX
DT 09-NOV-1998 (first entry)
XX
DE Mycobacterium tuberculosis antigen ESAT-6.
XX Tuberculosis; infection; diagnosis; antigen; ESAT-6.
KW
XX
OS Mycobacterium tuberculosis.
XX
PN WO9816645-A2.
PD
PD 23-APR-1998.
XX
PF 07-OCT-1997; 97WO-US18214.
XX
PR 13-MAR-1997; 97US-0818111.
PR 11-OCT-1996; 96US-0729622.
XX
PA (CORI-) CORIXA CORP.
XX
PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;
PI Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
XX
DR WPI; 1998-251292/22.
DR N-PSDB; AAV44393.
XX
PT New isolated Mycobacterium tuberculosis polypeptides and DNA - used
PT to develop products for the detection of M. tuberculosis infection
PT and diagnosis of tuberculosis
XX
PS Disclosure; Page 130; 250pp; English.
XX
CC This polypeptide comprises an antigenic portion of Mycobacterium
CC tuberculosis antigen ESAT-16. The invention relates to
CC compositions and methods for diagnosing tuberculosis. It provides
CC polypeptides (see AAW64291-W64379) comprising an antigenic portion
CC of a soluble M. tuberculosis antigen, or an immunogenic portion of
CC an M. tuberculosis antigen, as well as fusion proteins between
CC these polypeptides and known antigens such as ESAT-6. Also claimed
CC are methods and diagnostic kits for detecting M. tuberculosis
CC infection in a patient using these polypeptides, antibodies, or
CC oligonucleotide probes and primers.
XX
SQ Sequence 51 AA;

Query Match 54.1%; Score 259; DB 19; Length 51;
Best Local Similarity 100.0%; Pred. No. 9.3e-21;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEQOWNFAGIEAAASAIQGNVTSHSLDEGKQSLTKLAAAWGGSGSEAY 51
Db 1 MTEQOWNFAGIEAAASAIQGNVTSHSLDEGKQSLTKLAAAWGGSGSEAY 51

RESULT 16
AAV39131
ID AAV39131 standard; Protein; 51 AA.
XX
AC AAV39131;
XX
DT 05-NOV-1999 (first entry)

XX DE M. tuberculosis ESAT-6 amino acid sequence.
 XX
 KW Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;
 KW immunotherapy; diagnosis; immunisation; vaccine; infection;
 KW immune response; skin test.
 XX
 CS Mycobacterium tuberculosis.
 XX
 PN WO9942076-A2.
 XX
 ED 26-AUG-1999.
 XX
 FF 17-FEB-1999; 99WO-US03268.
 XX
 FF 05-MAY-1998; 98US-0072967.
 PR
 PR 18-FEB-1998; 98US-0025197.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
 PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
 PI
 XX WPI: 1999-527409/44.
 DR
 DR N-PSDB; AAZ19303.
 XX
 PT New antigens from Mycobacterium tuberculosis useful in diagnostic
 PT skin tests and protective or therapeutic vaccines or compositions
 XX
 PS Disclosure; Page 120; 299pp; English.
 XX
 CC The present invention describes polypeptides comprising an immunogenic
 CC part of a Mycobacterium tuberculosis antigen (Ag). Also described
 CC are vaccines and fusion protein containing M. tuberculosis Ag's.
 CC M. tuberculosis Ag's, DNAs encoding them, derived fusion proteins and
 CC other polypeptides fragments, can be used in pharmaceutical compositions
 CC or vaccines to generate a protective or therapeutic immune response to
 CC M. tuberculosis and as reagents in skin tests for diagnosis of
 CC tuberculosis. Ag can induce proliferation of, or cytokine secretion
 CC by, T, B or natural killer cells and/or macrophages in
 CC tuberculosis-immune subjects. AAZ19249 to AAZ19460 and AAY39083 to
 CC AAY39225 are used in the exemplification of the present invention.
 XX
 SQ Sequence 51 AA;
 Query Match 54.1%; Score 259; DB 20; Length 51;
 Best Local Similarity 100.0%; Pred. No. 9.3e-21;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MTEQQWNFAGIEAASAIQGNVTSIHSLDEGKSLTKLAAAWGSGSEAY 51
 Db 1 MTEQQWNFAGIEAASAIQGNVTSIHSLDEGKSLTKLAAAWGSGSEAY 51
 RESULT 17
 AAY38988
 ID AAY38988 standard; Protein; 51 AA.
 XX
 AC AAY38988;
 XX
 DT 05-NOV-1999 (first entry)
 XX
 XX M. tuberculosis recombinant antigen protein ESAT-6.
 DE
 DE Antigen; diagnosis; detection; infection; antibody; immunisation;
 KW vaccine; immunity.
 KW
 XX Mycobacterium tuberculosis.
 OS
 XX WO9942118-A2.
 PN
 PD 26-AUG-1999.
 XX

PF 17-FEB-1999; 99WO-US03265.
 XX
 PR 05-MAY-1998; 98US-0072596.
 PR
 PR 18-FEB-1998; 98US-0024753.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 XX Campos-Neto A, Dillon DC, Hendrickson RC, Houghton R;
 PI Lodes MJ, Reed SG, Skeiky YAW, Twardzik DR, Vedvick TS;
 PI
 XX WPI: 1999-527416/44.
 DR
 DR N-PSDB; AAZ19091.
 XX
 PT New polypeptide comprising antigenic portions of M. tuberculosis
 PT Claim 51; Page 165; 323pp; English.
 PS
 PS This invention describes novel recombinant antigens and their encoding
 CC nucleic acids derived from Mycobacterium tuberculosis. The novel
 CC polypeptides are useful for detecting M. tuberculosis infection in a
 CC biological sample by detecting antibodies which bind with the
 CC polypeptides, and are useful as vaccines for immunizing against
 CC M. tuberculosis infection. The new detection methods are needed as
 CC current vaccination strategies do not provide 100% immunity.
 XX
 SQ Sequence 51 AA;
 Query Match 54.1%; Score 259; DB 20; Length 51;
 Best Local Similarity 100.0%; Pred. No. 9.3e-21;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 MTEQQWNFAGIEAASAIQGNVTSIHSLDEGKSLTKLAAAWGSGSEAY 51
 Db 1 MTEQQWNFAGIEAASAIQGNVTSIHSLDEGKSLTKLAAAWGSGSEAY 51
 RESULT 18
 AAU01897
 ID AAU01897 standard; Protein; 51 AA.
 XX
 AC AAU01897;
 XX
 DT 29-AUG-2001 (first entry)
 XX
 DE Mycobacterium tuberculosis partial antigen ESAT-6.
 DE
 KW ESAT-6; antigen; vaccine; tuberculosis; AIDS;
 KW acquired immunodeficiency disease.
 KW
 XX Mycobacterium tuberculosis.
 OS
 XX WO200124820-A1.
 PN
 PD 12-APR-2001.
 XX
 XX 10-OCT-2000; 2000WO-US28095.
 PF
 PR 07-OCT-1999; 99US-0158338.
 PR
 PR 07-OCT-1999; 99US-0158425.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 XX Skeiky Y, Reed S, Houghton RL, McNeill PD, Dillon DC, Lodes ML;
 PI
 PI WPI: 2001-290576/30.
 DR
 DR N-PSDB; AAS03787.
 XX
 XX Vaccinating against Mycobacteria infections in mammals using fusion
 PT proteins comprising combinations of heterologous antigens -
 PT
 XX Disclosure; Page 164; 168pp; English.
 PS
 XX The sequence represents Mycobacterium tuberculosis ESAT-6 antigen.
 CC

Fri Feb 6 13:54:43 2004

us-09-805-427a-1.116min.rag

CC Compositions comprising at least 2 heterologous antigens, as a fusion
 CC protein, and vectors expressing the fusion proteins are used as vaccines
 CC to prophylactically immunise mammals (especially humans) against
 CC infection by Mycobacteria. The compositions contain at least 2
 CC heterologous antigens that increase the serological sensitivity of
 CC individuals infected with tuberculosis, a disease frequently affecting
 CC patients with acquired immunodeficiency disease, AIDS.
 XX
 SQ Sequence 51 AA;
 Query Match 54.1%; Score 259; DB 22; Length 51;
 Best Local Similarity 100.0%; Pred. No. 9.3e-21;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTEQWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAY 51
 DB 1 MTEQWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAY 51
 RESULT 19
 AAE29716
 ID AAE29716 standard; Protein; 51 AA.
 XX
 AC AAE29716;
 XX
 DT 27-JAN-2003 (first entry)
 XX
 DE Mycobacterium tuberculosis ESAT-6 antigenic protein.
 XX
 KW Vaccine; immunity; diagnostic agent; gene therapy; ESAT-6 antigen.
 XX
 OS Mycobacterium tuberculosis.
 XX
 PN WO200272792-A2.
 XX
 PD 19-SEP-2002.
 XX
 PF 13-MAR-2002; 2002WO-US08223.
 XX
 PR 13-MAR-2001; 2001US-275837P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky Y, Brannon M, Guderian J;
 XX
 DR WPI; 2002-759844/82.
 DR N-PSDB; AAD29716.
 XX
 PT New recombinant nucleic acid molecule comprising a Leishmania TSA,
 PT LeIF, M15 or 6H polynucleotide, useful as vaccine to elicit protective
 PT immunity against pathogenic microorganisms e.g. Leishmania and
 PT Mycobacterium tuberculosis
 XX
 PS Disclosure; Page 106; 155pp; English.
 XX
 CC The invention relates to a recombinant nucleic acid molecule encoding a
 CC fusion polypeptide. The recombinant nucleic acid comprises a heterologous
 CC polynucleotide sequence encoding an antigen or an antigenic fragment from
 CC Mycobacterium sp. and a Leishmania polynucleotide sequence encoding a
 CC polypeptide or its fragment. The Leishmania polynucleotide is selected
 CC from TSA, LeIF, M15, and 6H polynucleotides. Sequences of the invention
 CC are used in methods for eliciting immune response in mammals. They are
 CC useful as vaccines to elicit protective immunity against pathogenic
 CC microorganisms such as Leishmania and Mycobacterium tuberculosis. Fusion
 CC polypeptides are used for enhancing the expression of polynucleotides,
 CC as in vivo diagnostic agents and for raising antibodies in a non-human
 CC animal. The invention is used in gene therapy. The present sequence is
 CC M. tuberculosis ESAT-6 antigenic protein.
 XX
 SQ Sequence 51 AA;
 Query Match 54.1%; Score 259; DB 23; Length 51;
 Best Local Similarity 100.0%; Pred. No. 9.3e-21;

Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTEQWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAY 51
 DB 1 MTEQWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAY 51
 RESULT 20
 AAE17580
 ID AAE17580 standard; Protein; 51 AA.
 XX
 AC AAE17580;
 XX
 DT 22-APR-2002 (first entry)
 XX
 DE Mycobacterium species ESAT-6 protein.
 XX
 KW Fusion protein; antigen; serological sensitivity; immune response;
 KW tuberculosis; infection; vaccine; ESAT-6 protein.
 XX
 OS Mycobacterium sp.
 XX
 PN WO200198460-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 20-JUN-2001; 2001WO-US19959.
 XX
 PR 20-JUN-2000; 2000US-0597796.
 PR 01-FEB-2001; 2001US-265737P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skeiky Y, Reed S, Alderson M;
 XX
 DR WPI; 2002-147798/19.
 DR N-PSDB; AAD28351.
 XX
 PT Composition comprising MPT39 antigen and MPT32A antigen from
 PT Mycobacterium species, useful for eliciting immune response in a
 PT subject
 XX
 PS Claim 9; Page 122; 136pp; English.
 XX
 CC The present invention relates to fusion proteins containing at least
 CC two Mycobacterium species antigens, nucleotides encoding them and
 CC compositions comprising such fusion proteins. The present invention
 CC particularly relates to nucleic acids encoding fusion proteins that
 CC include two or more individual M. tuberculosis antigens which increase
 CC the serological sensitivity of sera from individuals infected with
 CC tuberculosis and methods for their use in diagnosis, prevention and
 CC treatment of tuberculosis infection. Sequences of the invention are
 CC useful for eliciting an immune response in a mammal, e.g., human,
 CC immunised with BCG. They are useful in the diagnosis, treatment and
 CC prevention of Mycobacterium infection. The fusion proteins and the
 CC polynucleotides are useful as diagnostic tools in patients infected
 CC with Mycobacterium, in vitro and in vivo assays for detecting humoral
 CC antibodies or cell-mediated immunity against M. tuberculosis, for the
 CC diagnosis of an infection or monitoring of disease progression, as
 CC immunogens to generate or elicit a protective immune response in a
 CC patient and for raising anti-M. tuberculosis antibodies in a non-human
 CC animal. Sequences of the invention are also used as vaccines. MPT32A
 CC fusion proteins of the invention are useful as in vivo diagnostic agents
 CC for intradermal skin test. The present sequence is Mycobacterium species
 CC ESAT-6 protein.
 XX
 SQ Sequence 51 AA;
 Query Match 54.1%; Score 259; DB 23; Length 51;
 Best Local Similarity 100.0%; Pred. No. 9.3e-21;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTEQWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAWGGSGSEAY 51

AAG90391
 ID AAG90391 standard: protein: ac aa

DT 22-APR-1998 (first entry)
 XX ESAT-6 antigen SEQ ID NO:119 from WO9738011.
 XX T-cell stimulatory peptide; immunogen; non-dendritic; carrier; tumour;
 KW scaffold; inhibition; metastasis; wound healing; solid phase.
 KW Synthetic.
 OS
 XX WO9738011-A1.
 XX 16-OCT-1997.
 XX 03-APR-1997; 97WO-DK00146.
 XX 03-APR-1996; 96DK-0000398.
 XX (PEPR-) PEPRESEARCH AS.
 XX Heegaard PMH, Jakobsen PH;
 XX WPI; 1997-512645/47.
 XX Non-dendritic peptide carrier linked to a solid phase - useful as a
 PT diagnostic agent and as a scaffold for production of chemical
 PT derivatives
 XX
 PS Example 32; Page 156; 262pp; English.
 XX
 CC A non-dendritic peptide carrier (A) has been developed which is coupled
 CC through a linker to a solid phase, forming a complex of (A)-solid phase.
 CC Where (A) comprises 10-50 amino acids capable of forming a secondary, and
 CC structure in a benign buffer after liberation from the solid phase, and
 CC further the (A)-solid phase complex comprises an immunogenic substance
 CC and/or an immune mediator coupled on (A). The present sequence
 CC represents a peptide used in an example from the present invention. An
 CC (A)-solid phase complex can be used as a scaffold for the production of
 CC chemical derivatives, characterised by covalently attaching molecules at
 CC attachment points. Alternatively (A) is used as a scaffold-peptide for
 CC the incorporation into an immunostimulating complex (iscom) resulting in
 CC (A)-iscom complex which is used for the chemical coupling of antigenic
 CC substances in an aqueous solution by conjugation. (A) derivatised with
 CC one or more peptides having fibronectin-, laminin- or vitronectin-like
 CC binding activities can be used for the promotion of cell-attachment to
 CC plastic surfaces, in particular to inhibit tumour growth and metastasis,
 CC and for promotion of wound healing. Also a derivatised (A) can be used
 CC for the selection of specifically-binding aptamers or as a diagnostic
 CC agent. Such diagnostic-(A) molecules could be used to detect molecules
 CC derived from or indicative of pregnancy or of a disease, such as an
 CC infectious, autoimmune or cancerous disease.
 CC (Updated on 25-MAR-2003 to correct PF field.)
 XX
 SQ Sequence 20 AA;
 Query Match 22.1%; Score 106; DB 18; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.7e-05; Indels 0; Gaps 0;
 Matches 20; Conservative 0; Mismatches 0;
 QY 51 YGVGVQKWDATATELNALQ 70
 DB 1 YGVGVQKWDATATELNALQ 20
 RESULT 24
 AAY88580
 ID AAY88580 standard; peptide; 20 AA.
 XX
 AC AAY88580;
 XX
 DT 11-AUG-2000 (first entry)
 XX Antigenic C-terminal sequence of ESAT-6.
 DE
 XX

KW Ligand presenting assembly; early secreted antigen target 6; ESAT-6;
 KW bacteria; detect; diagnosis; allergen; cancer; vaccine; immune response;
 KW neurotropic factor; autoimmune-system related compound; LPA; fungi;
 XX parasite; cell-adhesion molecule.
 OS Mycobacterium tuberculosis.
 XX
 FH Key Location/Qualifiers
 FT Cross-links 20
 FT /note= "Gln at position 20 is linked via a linker to Gln
 FT at position 20 of an identical peptide, where
 FT the linker is COCH2-CH(NH-Lys)-CH2CO, and the
 FT Lys residue contained in the linker is also
 FT linked to another peptide (see AAY88579)"
 XX
 FN WO200018791-A1.
 XX
 XX 06-APR-2000.
 PD
 XX 29-SEP-1999; 99WO-DK00510.
 PF
 XX 29-SEP-1998; 98DK-0001233.
 ER
 XX (STAT-) STATENS SERUM INST.
 PA (HOLM/) HOLM A.
 PA
 XX Holm A, Jorgensen RM, Ostergaard S, Theisen M;
 PI WPI; 2000-303438/26.
 XX
 DR New ligand presenting assemblies useful for diagnosis, treatment and
 XX prevention of diseases caused by e.g. viruses, bacteria, toxins,
 PT allergens, autoimmune system-related compounds, cancer-related
 PT compounds, cell adhesion molecules
 XX
 PS Claim 34; Page 81; 100pp; English.
 XX
 CC This sequence represents an antigenic C-terminal sequence from the
 CC Mycobacterium tuberculosis early secreted antigenic target 6 (ESAT-6)
 CC protein. The peptide is presented on the new ligand presenting assembly
 CC of the invention. The invention relates to method for preparing ligand
 CC presenting assemblies (LPAs) comprising:
 CC (a) providing by solid phase synthesis, or fragment coupling, ligands
 CC comprising desired sequences (e.g. the present sequence), the ligands
 CC being attached to a solid phase;
 CC (b) if necessary, deprotecting any N-terminal amino groups while the
 CC ligands are still attached to the solid phase;
 CC (c) reacting the ligands having unprotected N-terminal amino groups with
 CC an achiral di-, tri- or tetracarboxylic acid, to provide a construct
 CC having a ring structure; and
 CC (d) cleaving the construct from the solid phase, to provide an LPA
 CC comprising ligands having free C-terminal groups. The LPAs can be used
 CC for raising an immune response in an animal. They can also be used in
 CC vaccines and for generating antibodies in an animal. Alternatively they
 CC can be used for the treatment, alleviation, detection, diagnosis, or
 CC prophylaxis of diseases caused by viruses, bacteria, toxins, allergens,
 CC autoimmune system-related compounds, cancer related compounds, cell
 CC adhesion molecules, neurotropic factors, fungi or parasites. Use of the
 CC method enables the preparation of very long ring systems interconnected
 CC by reaction with the achiral di-, tri- or tetracarboxylic acid. The ring
 CC structure formed between desired sequences further enables additional
 CC presentation of desired sequences and chemical moieties. The LPAs provide
 CC very flexible systems for polyfunctional constructs, and furthermore,
 CC products of high purity are obtained.
 XX
 SQ Sequence 20 AA;
 Query Match 22.1%; Score 106; DB 21; Length 20;
 Best Local Similarity 100.0%; Pred. No. 9.7e-05;
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 51 YGVGVQKWDATATELNALQ 70
 ||||||||||||||||||||

Db 1 YQGVQCKWDATATLNALQ 20

RESULT 25

AAE12275
ID AAE12275 standard; peptide; 20 AA.

XX AC AAE12275;

XX DT 18-DEC-2001 (first entry)

XX DE Mycobacterium tuberculosis (Mtb) peptide #60.

XX KW Mycobacterium tuberculosis; Mtb peptide; antibacterial; vaccine;
XX KW infection; anti-Mtb immune response.

XX OS Mycobacterium tuberculosis.

XX FN WO200170774-A2.

XX PD 27-SEP-2001.

XX PF 20-MAR-2001; 2001WO-US08906.

XX PR 20-MAR-2000; 2000US-190834P.

XX PA (UYBR-) UNIV BROWN RES FOUND.

XX PL Degroot AS;

XX DR WPI; 2001-616401/71.

XX PT New vaccine for immunizing a mammalian subject, preferably humans,
XX PT against infection caused by Mycobacterium tuberculosis -

XX PS Disclosure; Fig 4; 42pp; English.

XX CC The present invention relates to Mycobacterium tuberculosis (Mtb)
XX CC vaccine candidate peptides. The invention also relates to a method
XX CC for identifying Mtb vaccine candidate peptides as well as vaccines
XX CC comprising these candidate peptides. Vaccines of the invention and
XX CC Mtb vaccine candidate peptides are useful for inducing an anti-
XX CC Mycobacterium tuberculosis (anti-Mtb) immune response by raising
XX CC anti-Mtb antibody in a mammalian subject preferably human. They
XX CC are used for immunising a mammalian subject, preferably humans,
XX CC against infection caused by Mycobacterium tuberculosis. The present
XX CC sequence is a Mtb vaccine candidate peptide.

XX SQ Sequence 20 AA;

Query Match 21.1%; Score 101; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00034;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 5 QWNFAGIEAASAIQGNVTS 24

Db 1 QWNFAGIEAASAIQGNVTS 20

RESULT 26

AAE12277
ID AAE12277 standard; peptide; 20 AA.

XX AC AAE12277;

XX DT 18-DEC-2001 (first entry)

XX DE Mycobacterium tuberculosis (Mtb) peptide #62.

XX KW Mycobacterium tuberculosis; Mtb peptide; antibacterial; vaccine;
XX KW infection; anti-Mtb immune response.

XX OS Mycobacterium tuberculosis.

XX PN WO200170774-A2.
XX PD 27-SEP-2001.
XX PF 20-MAR-2001; 2001WO-US08906.
XX PR 20-MAR-2000; 2000US-190834P.
XX PA (UYBR-) UNIV BROWN RES FOUND.
XX PI Degroot AS;
XX DR WPI; 2001-616401/71.
XX PT New vaccine for immunizing a mammalian subject, preferably humans,
XX PT against infection caused by Mycobacterium tuberculosis -
XX PS Disclosure; Fig 4; 42pp; English.
XX CC The present invention relates to Mycobacterium tuberculosis (Mtb)
XX CC vaccine candidate peptides. The invention also relates to a method
XX CC for identifying Mtb vaccine candidate peptides as well as vaccines
XX CC comprising these candidate peptides. Vaccines of the invention and
XX CC Mtb vaccine candidate peptides are useful for inducing an anti-
XX CC Mycobacterium tuberculosis (anti-Mtb) immune response by raising
XX CC anti-Mtb antibody in a mammalian subject preferably human. They
XX CC are used for immunising a mammalian subject, preferably humans,
XX CC against infection caused by Mycobacterium tuberculosis. The present
XX CC sequence is a Mtb vaccine candidate peptide.

SQ Sequence 20 AA;

Query Match 20.5%; Score 98; DB 22; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00071;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 18 IQGNVTSIHSLDEGQSILT 37

Db 1 IQGNVTSIHSLDEGQSILT 20

RESULT 27

AAAY21964
ID AAY21964 standard; Protein; 96 AA.

XX AC AAY21964;

XX DT 06-SEP-1999 (first entry)

XX DE Amino acid sequence of antigen ORF7-1.

XX KW Immunogenic; Mycobacterium tuberculosis; immune response; infection;
XX KW tuberculosis; fusion polypeptide; T-cell epitope; ESAT-6; MPT59; TB;
XX KW pharmaceutical; vaccination; M. africanum; M. bovis; CFP7A; CFP30A;
XX KW CFP7B; CFP19; CFP27; CFP30A; RDI-ORF; CFP10A; CFP16; CFP19; CFP23;
XX KW CFP25A; CFP30B; CFP7B.

XX OS Mycobacterium tuberculosis.

XX PN WO9924577-A1.

XX PD 20-MAY-1999.

XX PF 08-OCT-1998; 98WO-DK00438.

XX PR 01-APR-1998; 98WO-DK00132.

XX PR 10-NOV-1997; 97DK-0001277.

XX PR 05-JAN-1998; 98US-0070486.

XX PA (STAT-) STATENS SERUM INST.

XX PI Andersen P, Skjot R;

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:35:17 : Search time 6.3333 Seconds
(without alignments)
1442.530 Million cell updates/sec

Title: US-09-805-427A-1
Perfect score: 479
Sequence: 1 MTEQQWNFAGIEAASALQG.....ISEAGQMASTEGNVTGMFA 95

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283102

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database :

PIR 76:*

1: Pirl.*

2: Pirl2.*

3: Pirl3.*

4: Pirl4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	100.0	95	2 A70803	early secretory an
2	186	38.8	95	2 T10031	early secretory an
3	96	20.0	96	2 F70857	hypothetical prote
4	86.5	18.1	100	2 B70977	hypothetical prote
5	85	17.7	96	2 F70836	hypothetical prote
6	81	16.9	96	2 H87225	hypothetical prote
7	80.5	16.8	90	2 A70600	probable cell suri
8	80	16.7	498	1 FLEC	hypothetical prote
9	79	16.5	104	2 C86954	flagellin - Escher
10	77.5	16.2	125	2 C70977	probable secreted
11	76	15.9	1037	2 D91045	hypothetical prote
12	75	15.7	490	2 H83286	aminoglycoside eff
13	75	15.7	1037	2 E65022	probable chemotaxi
14	75	15.7	1037	2 G85889	acriflavin resista
15	73.5	15.3	371	2 G96937	hypothetical prote
16	73	15.2	1204	2 B83566	cell wall-associat
17	72.5	15.1	238	2 T05959	probable bacteriop
18	72.5	15.1	756	2 A43582	1,3-beta-glucanase
19	72	15.0	96	2 D83771	surface antigen ms
20	72	15.0	637	2 C87322	hypothetical prote
21	72	15.0	1225	2 F75605	methyl-accepting c
22	72	15.0	1697	2 T00079	minor tail protein
23	72	15.0	6713	2 B89921	hypothetical prote
24	71.5	14.9	94	2 D70608	hypothetical prote
25	71.5	14.9	526	2 D75391	AlpP-related prote
26	71.5	14.9	927	2 A91739	transmembrane prot
27	71	14.8	97	2 H96904	uncharacterized sm
28	71	14.8	356	2 S02017	regulatory protein
29	70.5	14.7	94	2 B70930	hypothetical prote

30	70.5	14.7	94	2 C70662	hypothetical prote
31	70.5	14.7	94	2 D70560	hypothetical prote
32	70.5	14.7	388	2 A46173	Mip4 protein - Str
33	70.5	14.7	656	2 A72428	methyl-accepting c
34	70.5	14.7	656	2 E72379	methyl-accepting c
35	70.5	14.7	661	2 G72316	methyl-accepting c
36	70.5	14.7	1510	2 T13634	probable minor tai
37	70.5	14.7	2073	2 T43311	fatty acyl-CoA syn
38	70.5	14.7	2073	2 T39207	fatty acid synthas
39	70.5	14.7	2475	2 T00047	gellan lyase (EC 4
40	70	14.6	103	2 B70600	hypothetical prote
41	70	14.6	649	2 T45102	H+-transporting tw
42	70	14.6	710	1 I51283	hepatocyte growth
43	69.5	14.5	382	2 S15578	ipac protein - Shi
44	69.5	14.5	1509	2 B89985	hypothetical prote
45	69	14.4	572	2 C96024	probable glycerone
46	69	14.4	1037	2 A80816	probable efflux pu
47	68.5	14.3	103	2 T10920	3C3.11 protein - S
48	68	14.2	164	2 A45512	cold-regulated pro
49	68	14.2	335	2 T02211	1,3-beta-glucanase
50	68	14.2	388	2 S52536	fcra 15 protein -
51	68	14.2	405	2 A33939	Fc gamma (IgG) rec
52	68	14.2	409	2 S10485	licheninase (EC 3.
53	68	14.2	701	2 UC7184	gamma adducin - hu
54	67.5	14.1	238	2 T05962	flagellin - Shigel
55	67.5	14.1	364	2 S78461	phosphoglycerate m
56	67.5	14.1	395	2 D83782	hypothetical prote
57	67.5	14.1	495	2 S60175	flagellin - Shigel
58	67.5	14.1	524	2 S44982	regulatory protein
59	67	14.0	231	2 AF1503	flagellin - Shigel
60	67	14.0	293	2 D96560	hypothetical prote
61	67	14.0	451	2 B70805	hypothetical prote
62	67	14.0	550	2 S44980	flagellin - Shigel
63	67	14.0	562	2 D85439	tubulin-like prote
64	67	14.0	565	2 T41061	flagellin - Escher
65	67	14.0	1063	2 A40253	acidic nuclear pro
66	67	14.0	1738	1 A24558	complement C4 prec
67	67	14.0	2094	2 S33124	tpa protein - huma
68	66.5	13.9	3635	2 T10053	laminin alpha 5 ch
69	66.5	13.9	288	2 A99305	hypothetical prote
70	66.5	13.9	428	1 D71155	probable methyl-ac
71	66.5	13.9	454	2 A41659	benzoate 1,2-dioxy
72	66.5	13.9	457	2 G82925	hypothetical prote
73	66.5	13.9	595	2 B48658	flagellin - Escher
74	66.5	13.9	706	2 G69371	acetyl-CoA synthet
75	66.5	13.9	709	2 S75212	methyl-accepting c
76	66.5	13.9	709	2 S75212	comE protein - Syn
77	66.5	13.9	1819	2 D97132	uncharacterized ph
78	66	13.8	424	2 S72695	L-aspartate oxidas
79	66	13.8	455	2 JC7550	glucokinase (EC 2.
80	66	13.8	527	2 D87062	L-aspartate oxidas
81	66	13.8	530	2 S62439	hypothetical serin
82	66	13.8	626	2 F85295	hypothetical prote
83	66	13.8	627	2 T05789	hypothetical prote
84	66	13.8	927	2 AH1369	transmembrane prot
85	66	13.8	1467	2 A75564	conserved hypoteth
86	65.5	13.7	137	2 JH0127	mobB protein - Esc
87	65.5	13.7	465	2 C43357	triacylglycerol li
88	65.5	13.7	531	2 D83297	probable chemotaxi
89	65.5	13.7	585	2 F90961	flagellin importe
90	65.5	13.7	585	2 F85809	hypothetical prote

ALIGNMENTS

RESULT 1

A70803

early secretory antigen target esat6 - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence.revision 17-Jul-1998 #text_change 22-Oct-1999

C/Accession: A70803; S49174

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70803
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-95 <COL>
A:Cross-references: GB:AL021220; GB:AL123456; NID:g3261558; PIDN:CAA17967.1; PID:el26412
A:Experimental source: strain H37RV
R:Soerensen, A.L.; Nagai, S.; Houn, G.; Andersen, P.; Andersen, A.B.
submitted to the EMBL Data Library, June 1994
A:Reference number: S49174
A:Accession: S49174
A:Molecule type: DNA
A:Residues: 1-13, 'R', 15-22, 'S', 24-95 <SOE>
A:Cross-references: EMBL:X79562; NID:g531708
C:Genetics:
A:Gene: csat6

Query Match 100.0%; Score 479; DB 2; Length 95;
Best Local Similarity 100.0%; Pred. No. 1.3e-39;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEGQSILTKLAAAGGSGSEAYQGVQKWDA 60
DB 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEGQSILTKLAAAGGSGSEAYQGVQKWDA 60

QY 61 TATELNALONLARTISEAGQAMASTEGNVTGMFA 95
DB 61 TATELNALONLARTISEAGQAMASTEGNVTGMFA 95

RESULT 2
T10031
early secretory antigen target 6 protein homolog - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 21-Jul-2000
C:Accession: T10031
R:Eiglmeyer, K.; Honore, N.; Woods, S.A.; Caudron, B.; Cole, S.T.
Mol. Microbiol. 7, 197-206, 1993
A:Title: Use of an ordered cosmid library to deduce the genomic organization of Mycobact
A:Reference number: Z16917; MUID:93188700; PMID:8446027
A:Accession: T10031
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-95 <EIG>
A:Cross-references: EMBL:Y14967; NID:g2370268; PIDN:CAA75200.1; PID:g2370279
C:Genetics:
A:Gene: esx

Query Match 38.8%; Score 186; DB 2; Length 95;
Best Local Similarity 36.3%; Pred. No. 3.3e-11;
Matches 33; Conservative 27; Mismatches 31; Indels 0; Gaps 0;

QY 4 QOWNFAGIEAASAIQGNVTSIHSLDEGQSILTKLAAAGGSGSEAYQGVQKWDA 63
DB 3 QAWHFPALQAVNELQGSQSRIDALLEQCESLTKLQSSWHGSGNESYVQRFNQNT 62

QY 64 ELNNALONLARTISEAGQAMASTEGNVTGMF 94
DB 63 GINHALGDLVQAINHSAFTMQTEAGVMSMF 93

RESULT 3
F70857
hypothetical protein Rv3019c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
A:Accession: F70857
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Nature 393, 537-544, 1998

; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: F70857
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-96 <COL>
A:Cross-references: GB:AL021287; GB:AL123456; NID:g3261508; PIDN:CAA16104.1; PID:el23777
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv3019c

Query Match 20.0%; Score 96; DB 2; Length 96;
Best Local Similarity 30.2%; Pred. No. 0.018;
Matches 29; Conservative 16; Mismatches 39; Indels 12; Gaps 3;

QY 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEGQSILTKLAAAGGSGSEAYQGVQKWDA 60
DB 1 MSQIMNYNPAMMAHAGDWAGVAGTQSLGADIASQAVLSSAQWGDGTGTYQGWQW-- 58

QY 61 TATELNALONLARTISEAGQAMASTEGNVTGMFA 95
DB 59 -----NOALEDLVR-----AYQSMGTHESNTMAMLA 85

RESULT 4
B70977
hypothetical protein Rv3444c - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: B70977
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: B70977
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-100 <COL>
A:Cross-references: GB:Z95389; GB:AL123456; NID:g3242261; PIDN:CA808674.1; PID:es16051
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv3444c

Query Match 18.1%; Score 86.5; DB 2; Length 100;
Best Local Similarity 24.2%; Pred. No. 0.15;
Matches 22; Conservative 13; Mismatches 55; Indels 1; Gaps 1;

QY 6 MNFAGIEAASAIQGNVTSIHSLDEGQSILTKLAAAGGSGSEAYQGVQKWDA 64
DB 9 YNFDALEYSVRQEIHTTAARFNAALQELRSQAPLQQLWTREAAAYHAEQLKWHQAASA 68

QY 65 LNNAALONLARTISEAGQAMASTEGNVTGMFA 95
DB 69 LNEILLDLGNVVRHGADDDVAHADRRRAAGAWA 99

RESULT 5
F70836
hypothetical protein Rv0288 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
A:Accession: F70836
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A;Residues: 1-15,140-402,489-498 <KU2>
A;Cross-references: GB:M21445; NID:G146313; PIDN:AAA23951.1; PID:G146314
C;Comment: This is the principal protein component of the flagellar filament, which is a
C;Genetics:
A;Gene: fliC; hag
A;Map position: 42 min
C;Superfamily: flagellin
C;Keywords: flagellum

Query Match 16.7%; Score 80; DB 1; Length 498;
Best Local Similarity 30.4%; Pred. No. 4.3;
Matches 24; Conservative 9; Mismatches 38; Indels 8; Gaps 1;

QY 9 AGIEAASAIQGNVTSHSLDEGKSLTKLAAANGSGSEAYGVQVKWDATATLNA 68

DB 394 AGCEALTAVANGKTDPLKALDDATASVDKFRSLG-----AVQRLDSAVTNLNT 445

QY 69 LONLARTISEAQMASTE 87

DB 446 TTNLSEAQRIOADYATE 464

RESULT 9

C86954 Probable secreted protein [imported] - Mycobacterium leprae

C;Species: Mycobacterium leprae

C;Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C;Accession: C86954

R;Cole, S.T.; Eiglmeyer, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho

R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,

ean, M.A.; Rutherford, K.M.

Nature 409, 1007-1011, 2001

A;Title: Massive gene decay in the leprosy bacillus.

A;Reference number: A86909; MUID:21128732; PMID:11234002

A;Accession: C86954

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-104 <STO>

A;Cross-references: GB:AL450380; NID:G13092644; PIDN:CAC29871.1; GSPDB:GN00147

C;Genetics:

A;Gene: ML0363

Query Match 16.5%; Score 79; DB 2; Length 104;

Best Local Similarity 25.8%; Pred. No. 0.86;

Matches 24; Conservative 14; Mismatches 50; Indels 6; Gaps 1;

QY 2 TEQWNFAGTEAASAIQGNVTSHSLDEGKSLTKLAAANGSGSEAYGVQVKWDAT 61

DB 16 TKSNTPLAGKSTARPA-----RFNALELRSQIAPLQQRWTRAAVVVPVQLKWHQA 69

QY 62 ATELNALQNLARTISEAQMASTEIGNVTGMPA 95

DB 70 ATALNEILDVGNVIRAGAEVANTDRSGGWA 103

RESULT 10

C70977

hypothetical protein Rv3445c - Mycobacterium tuberculosis (strain H37RV)

C;Species: Mycobacterium tuberculosis

C;Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

C;Accession: C70977

R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.

Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A;Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A;Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A;Reference number: A70500; MUID:98295987; PMID:9634230

A;Accession: C70977

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-125 <COL>

A;Cross-references: GB:Z95389; GB:AL123456; NID:G3242261; PIDN:CAE08675.1; PID:e316052;
A;Experimental source: strain H37RV
C;Genetics:
A;Gene: RV3445c

Query Match 16.2%; Score 77.5; DB 2; Length 125;

Best Local Similarity 35.3%; Pred. No. 1.5;

Matches 18; Conservative 8; Mismatches 18; Indels 7; Gaps 1;

QY 41 AAWGSGSEAYGVQVKWDATATLNNALQNLARTI-----SEAGAMA 84

DB 65 SYWGGLAAARFQVDVVRWNAESTRLYHVLHATATIRNEAALRAGQIHA 115

RESULT 11

D91045

aminoglycoside efflux pump [imported] - Escherichia coli (strain O157:H7, substrain RIM

C;Species: Escherichia coli

C;Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001

C;Accession: D91045

R;Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.C.

gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.

DNA Res. 8, 11-22, 2001

A;Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gen

A;Reference number: A99629; MUID:21156231; PMID:11258796

A;Accession: D91045

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1037 <HAY>

A;Cross-references: GB:BA000007; PIDN:BAB36755.1; PID:G13362802; GSPDB:GN00154

A;Experimental source: strain O157:H7, substrain RIMD 0509952

C;Genetics:

A;Gene: EC83332

C;Superfamily: acriflavin resistance protein

Query Match 15.9%; Score 76; DB 2; Length 1037;

Best Local Similarity 26.2%; Pred. No. 25;

Matches 27; Conservative 18; Mismatches 38; Indels 20; Gaps 4;

QY 7 NFAGIEA-----AASAIQGNVTSHSL-----DEGKSLTKLAAANGSGSEAYQG 53

DB 48 NYPGSAQTLENTVTQVIEQNMNTGLDNLMTMSSQSGTQASVT-LSFKAGTDPDEAVQQ 106

QY 54 VQKWDATATLNNALQNLARTISEAQ-----AMASTEIGNV 90

DB 107 VQNLQSAQMKLPQAVQNGQLTVRKTGDTNLTIAFVSTDGSM 149

RESULT 12

H83286

Probable chemotaxis transducer PA2867 [imported] - Pseudomonas aeruginosa (strain PA01

C;Species: Pseudomonas aeruginosa

C;Date: --5-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C;Accession: H83286

R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Li

; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat

A;Reference number: AB2950; MUID:20437337; PMID:10984043

A;Accession: H83286

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-490 <STO>

A;Cross-references: GB:AE004713; GB:AE004091; NID:G9948952; PIDN:AAG06255.1; GSPDB:GN0

A;Experimental source: strain PA01

C;Genetics:

A;Gene: PA2867

Query Match 15.7%; Score 75; DB 2; Length 490;

Best Local Similarity 29.1%; Pred. No. 13;

Matches 23; Conservative 12; Mismatches 36; Indels 8; Gaps 2;


```
QY      8  FAGTEAARASAIQGNVTSIHSLDEGKQSLTKLAAAGWGSSEAYGQVQKWDATATELNN 67
      201 FNAQAGYQRIYGTVAATAATKLDGAQALARM-----GQVRQGMGLGQOSETDQTATAIN- 255
      68 ALQNLARTISEAGQAMAST 86
      256 ---EMSTIVFHIAQHAUT 271

RESULT 13
E65022
acriflavin resistance protein acrD - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C:Date: 12-Sep-1997 #sequence_revision 17-Sep-1997 #text_change 01-Mar-2002
C:Accession: E65022; C42959; S26997
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;
  .A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: E65022
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1037 <BLAT>
A:Cross-references: GB:AE000334; GB:U00096; NID:g1788813; PIDN:AACT5523.1; PID:g1788814;
A:Experimental source: strain K-12, substrain MGL1655
R:Bouvier, J.; Richaud, C.; Higgins, W.; Bogler, O.; Stragier, P.
J. Bacteriol. 174, 5265-5271, 1992
A:Title: Cloning, characterization, and expression of the dapE gene of Escherichia coli.
A:Reference number: A42959; MUID:92355499; PMID:1644752
A:Accession: C42959
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 998-1037 <BOU>
A:Cross-references: GB:X57403; GB:S41761; NID:g41231; PIDN:CAA40663.1; PID:g41232
A:Note: sequence extracted from NCBI backbone (NCBIN:110333, NCBI:P:110334)
R:Stragier, P.
submitted to the EMBL Data Library, January 1991
A:Reference number: S26997
A:Accession: S26997
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 998-1037 <STP>
A:Cross-references: EMBL:X57403; NID:g41231; PIDN:CAA40663.1; PID:g41232
A:Experimental source: strain K-12
C:Genetics:
A:Gene: acrD
C:Superfamily: acriflavin resistance protein
C:Keywords: transmembrane protein
F;14-30/Domain: transmembrane #status predicted <TM1>
F;343-359/Domain: transmembrane #status predicted <TM2>
F;370-386/Domain: transmembrane #status predicted <TM3>
F;397-413/Domain: transmembrane #status predicted <TM4>
F;442-458/Domain: transmembrane #status predicted <TM5>
F;472-488/Domain: transmembrane #status predicted <TM6>
F;540-556/Domain: transmembrane #status predicted <TM7>
F;873-889/Domain: transmembrane #status predicted <TM8>
F;897-913/Domain: transmembrane #status predicted <TM9>
F;928-944/Domain: transmembrane #status predicted <TM10>
F;972-988/Domain: transmembrane #status predicted <TM11>
F;1009-1025/Domain: transmembrane #status predicted <TM12>

Query Match      15.7%; Score 75; DB 2; Length 1037;
Best Local Similarity 26.2%; Pred. No. 31;
Matches 27; Conservative 18; Mismatches 38; Indels 20; Gaps 4;

QY      7  NFAGIEA-----AASAIQGNVTSIHSL-----DEGKQSLTKLAAAGWGSSEAYGQ 53
      48  NYPGASAQTLENTVTQVIEQNMTGLDNLMTSSQSGTGQASVT-LSFKAGTDPEAVQQ 106
      54  VQKWDATATLNNALQNLARTISEAGQ-----AMASTEAGNV 90
      107  VQNLQSAMRKLPOAVQNGVTVRKGTDLNLTITAFVSTDGSM 149

RESULT 14
G85889
hypothetical protein acrD [imported] - Escherichia coli (strain O157:H7, substrain EDL93)
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: G85889
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
  iller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
  Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: G85889
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1037 <STO>
A:Cross-references: GB:AE005174; NID:g12516848; PIDN:AAAG57579.1; GSPDB:GN00145; UWGP:237
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: acrD
C:Superfamily: acriflavin resistance protein

Query Match      15.7%; Score 75; DB 2; Length 1037;
Best Local Similarity 26.2%; Pred. No. 31;
Matches 27; Conservative 18; Mismatches 38; Indels 20; Gaps 4;

QY      7  NFAGIEA-----AASAIQGNVTSIHSL-----DEGKQSLTKLAAAGWGSSEAYGQ 53
      48  NYPGASAQTLENTVTQVIEQNMTGLDNLMTSSQSGTGQASVT-LSFKAGTDPEAVQQ 106
      54  VQKWDATATLNNALQNLARTISEAGQ-----AMASTEAGNV 90
      107  VQNLQSAMRKLPOAVQNGVTVRKGTDLNLTITAFVSTDGSM 149

RESULT 15
G96937
cell wall-associated hydrolase [imported] - Clostridium acetobutylicum
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: G96937
R:Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
  .J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
  .
A:Reference number: A96900; MUID:21359325; PMID:21359325
A:Accession: G96937
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-371 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78290.1; PID:g15023152; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0309

Query Match      15.3%; Score 73.5; DB 2; Length 371;
Best Local Similarity 28.6%; Pred. No. 13;
Matches 24; Conservative 17; Mismatches 32; Indels 11; Gaps 3;

QY      10  GIEAASAIQGNVTSIHSLDEGKQSLTKLAAAGWGSSEAYGQVQKWDATAT---ELN 66
      13  GITVSCS---GNIVFASPLQDQYNQSQQQYQNA----LKSVDIENKIEALDNQIGELN 64
      67  NALQNLARTISEAGQAMASTEAGNV 90
      65  NSINDTDKRINESQNMAITQGI 88

RESULT 16
B83566
probable bacteriophage protein PAO641 [imported] - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
```

C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: B83566
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim, J.; Lory, S.; Olson, M.V.
A:Reference: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen. Nature 406, 959-964, 2000
A:Reference number: B83566
A:Accession: B83566
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1204 <STO>
A:Cross-references: GB:AE004499; GB:AE004091; NID:g9946508; PIDN:AAG04030.1; GSPDB:GN001
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0641

Query Match 15.2%; Score 73; DB 2; Length 1204;
Best Local Similarity 23.7%; Pred. No. 59;
Matches 23; Conservative 25; Mismatches 35; Indels 14; Gaps 4;

QY 11 TBAASATQGNVTSIHSLDEKQS-----LTKLAAWGGSGS--RAYQVQKWDATATE 64
DB 993 LEATADAGRTATVERVATSDQATQRLDLSASIGGTAASLOEQATARANADSALAQ 1052
QY 65 LNNALQNLART-----TSEAGQAMASTEGRVTGMFA 95
DB 1053 RDTVQ--ARTDNGAAQTTSQATSLDGNVKNYS 1087

RESULT 17
T05959
1.3-beta-glucanase (EC 3.2.1.-) Glu5 - soybean (fragment)
C:Species: Glycine max (soybean)
C>Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 29-Sep-1999
C:Accession: T05959
R:Jin, W.; Horner, H.T.; Palmer, R.G.; Shoemaker, R.C.
A:Title: Submitted to the EMBL Data Library, November 1997
A:Description: Soybean beta-1,3-glucanase.
A:Reference number: Z15471

A:Accession: T05959
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-238 <JIN>
A:Cross-references: EMBL:AF034110; NID:g2921319; PIDN:AAC04712.1; PID:g2921320
A:Experimental source: strain Minsoy
C:Genetics:
A:Gene: Glu5
C:Superfamily: beta-1,3-glucanase
C:Keywords: glycosidase; hydrolase; polysaccharide degradation

Query Match 15.1%; Score 72.5; DB 2; Length 238;
Best Local Similarity 30.9%; Pred. No. 9.7;
Matches 30; Conservative 15; Mismatches 35; Indels 17; Gaps 5;

QY 12 EAAASAIQGNVTSIHSLDEKQSILTKLAAA--WGGSGSEAY-QGVQKWDATATELN-- 66
DB 38 EGVQLAQRSGNTEVILGVNDQLQSLTNAGATNWNKVKYASQNVKXYIAGVNEIHP 97
QY 67 -----NALQNLARTISEA---GQAMASTEGRVT 91
DB 98 GDSLAGSVLPALENTQKASIAANLQGMKVSTADITD 134

RESULT 18
A43582
Surface antigen msp1 beta chain - *Anaplasma marginale*
C:Species: *Anaplasma marginale*
C>Date: 12-Jan-1993 #sequence_revision 12-Jan-1993 #text_change 08-Oct-1999
C:Accession: A43582
R:Barbet, A.F.; Allred, D.R.
A:Title: The msp1beta multigene family of *Anaplasma marginale*: nucleotide sequence analysis

A:Reference number: A43582; MUID:911147241; PMID:1671779
A:Accession: A43582
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-756 <BAR>
A:Cross-references: GB:M59845; NID:g142172; PIDN:AAA22066.1; PID:g142173
C:Keywords: surface antigen

Query Match 15.1%; Score 72.5; DB 2; Length 756;
Best Local Similarity 21.7%; Pred. No. 38;
Matches 18; Conservative 18; Mismatches 40; Indels 7; Gaps 1;

QY 8 FAGTEAAASAIQGNVTSIHSLDEKQSILTKLAAWGGSGSEAYQGVQKWDATATELN 67
DB 383 FGGINTCAKLMKGSFAALNQKFEINATLTREA-----TDIAQGVKETYQSIGDAFNC 435

QY 68 ALQNLARTISEAGQAMASTEGRNV 90
DB 436 AFKSVGDFAKSGIDAFKSGANDGI 458

RESULT 19
D83771
Hypothetical protein BH0972 [imported] - *Bacillus halodurans* (strain C-125)
C:Species: *Bacillus halodurans*
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 15-Jun-2001
C:Accession: D83771
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hara, Y.
A:Title: Complete genome sequence of the alkaliphilic bacterium *Bacillus halodurans* and its relationship to other members of the family Halobacteriaceae.
A:Reference number: A83650; MUID:20512582; PMID:11058132

A:Accession: D83771
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-96 <STO>
A:Cross-references: GB:AP001510; GB:BA000004; NID:g10173440; PIDN:BA804691.1; GSPDB:GN0
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH0972

Query Match 15.0%; Score 72; DB 2; Length 96;
Best Local Similarity 21.4%; Pred. No. 3.7;
Matches 21; Conservative 16; Mismatches 43; Indels 18; Gaps 3;

QY 1 MTEQWNPAGIEAAASAIQGNVTSIHSLDEKQSILTKLAAWGGSGSEAYQGVQKWD 60
DB 6 MTPPE-----LRTANETRSDAASYQDILLQRGDARIEKLGSTWEG---EAFQGFQAQWOD 57

QY 61 TATELNALQ-----NLARTISEAGQAMASTEGR 88
DB 58 KRKHVEELIQLYEEELGAQTDDIANVVTDDQEIERSIG 95

RESULT 20

C87322
methyl-accepting chemotaxis protein McpG [imported] - *Caulobacter crescentus*
C:Species: *Caulobacter crescentus*
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 10-May-2001
C:Accession: C87322
R:Niermar, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Kolic, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.; Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
A:Reference number: A87249; MUID:21173698; PMID:11259647

A:Accession: C87322
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-637 <STO>
A:Cross-references: GB:AE005673; NID:g13421789; PIDN:AAK22575.1; GSPDB:GN00148
C:Genetics:
A:Gene: CC0589
C:Superfamily: methyl-accepting chemotaxis protein mcpA

Query Match 15.0%; Score 72; DB 2; Length 637;
Best Local Similarity 25.5%; Pred. No. 35;
Matches 24; Conservative 16; Mismatches 42; Indels 12; Gaps 3;

QY 9 AGTEAASAIQGN-----VTSIHLLDEGKQSLTKLAAANGSGSEAYQGVQKWDATAT 63
DB 476 AGVEAARAGSGRGFAVVAQEVSLAQRSQAQSREITALVNSARADVAAQGGREV-SSTGA 534

QY 64 ELNN-----ALQNLARTISEAGQAMASTEGNVT 91
DB 535 ALNAIVSRVLEIHGLVERISASNEAQAQGTIGSIT 568

RESULT 21
F75605
minor tail protein gp26-related protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
C:Accession: F75605
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: F75605
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1225 <WHI>
A:Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12240.1; PID:g646053
A:Experimental source: strain R1
A:Gene: DR0108
A:Map position: 2

Query Match 15.0%; Score 72; DB 2; Length 1225;
Best Local Similarity 29.3%; Pred. No. 75;
Matches 27; Conservative 15; Mismatches 36; Indels 14; Gaps 4;

QY 9 AGTEAASAIQGNVTST-----HSLLDEGKQSLTKLAAANGSGSEAYQGVQKWDAT 61
DB 53 AGTIAAGLLATGATVAVVGSTVGLAQEAQAQVNDFOAKLGASREEA-----EKLGTV 107

QY 62 ATELNALQNLARTISEAGQAMASTEGNVTGM 93
DB 108 AEQVFG--DNWGTSLSEAGEAVANVRVKGL 137

RESULT 22
T00079
hypothetical protein KIAA0465 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 21-Jul-2000
C:Accession: T00079
R:Seiki, N.; Ohira, M.; Nagase, T.; Ishikawa, K.; Miyajima, N.; Nakajima, D.; Nomura, N.;
DNA Res. 4, 345-349, 1997
A:Title: Characterization of cDNA clones in size-fractionated cDNA libraries from human
A:Reference number: Z14085; MUID:98116662; PMID:9455484
A:Accession: T00079
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1697 <SEK>
A:Cross-references: EMBL:AB007934; NID:g3413891; PIDN:BAA32310.1; PID:g3413892
A:Experimental source: brain
C:Genetics:
A:Map position: 1
A:Note: KIAA0465

Query Match 15.0%; Score 72; DB 2; Length 1697;
Best Local Similarity 27.3%; Pred. No. 1.1e+02;
Matches 24; Conservative 14; Mismatches 28; Indels 22; Gaps 3;

QY 16 SAIQGNVTSIHSLLDGKQSLTKLAAANGSGSEAYQGV---QKWDAT----- 61
DB 689 SOLKAKEETYNQLLDKGR--LMLLSRDDSGSGSKTEQSVALLDQKWHVSSKMEERKSKL 746

QY 62 -----ATELNALQNLARTISEAGQAM 83
DB 747 EEALNLFATEFQNSLQBFINWLTIAEQSL 774

RESULT 23
B89921
hypothetical protein ebha [imported] - Staphylococcus aureus (strain N315)
C:Species: Staphylococcus aureus
C>Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 22-Oct-2001
C:Accession: B89921
R:Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguc
ma, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.;
C.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramatsu, K.
Lancet 357, 1225-1240, 2001
A:Title: Whole genome sequencing of methicillin-resistant *Staphylococcus aureus*.
A:Reference number: A89758; MUID:21311952; PMID:11418146
A:Accession: B89921
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-6713 <KUR>
A:Cross-references: GB:BA000018; PID:g13701232; PIDN:BAB42527.1; GSPDB:GN00149
A:Experimental source: strain N315
C:Genetics:
A:Gene: ebha

Query Match 15.0%; Score 72; DB 2; Length 6713;
Best Local Similarity 30.9%; Pred. No. 5.5e+02;
Matches 21; Conservative 15; Mismatches 26; Indels 6; Gaps 2;

QY 14 AASAIQGNVTSIHSLLDGKQSLTKLAAANGSGSEAYQGVQKWDATATLNALQNL 73
DB 1338 AKTTAKNNLNLTSINNAQKDALTR-----NIDGATTVAGVNOE-TAKATELNAMHSLQ 1391

QY 74 RTISEAQG 81
DB 1392 NGINDETQ 1399

RESULT 24
D70608
hypothetical protein Rv1198 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: D70608
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70608
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-94 <COL>
A:Cross-references: GB:Z93777; GB:AL123456; NID:g3261726; PIDN:CAB07821.1; PID:e311075;
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv1198

Query Match 14.9%; Score 71.5; DB 2; Length 94;
Best Local Similarity 25.5%; Pred. No. 4.1;
Matches 24; Conservative 20; Mismatches 31; Indels 19; Gaps 4;

QY 6 WNFAGTAAASAIQGNVTSIHSLLDGKQSLTK-----LAAA--WGGSGSEAYQGV----- 53
DB 5 YQFQVDVAHGAMTAAQA-----GLLEAEHQAIIRDVLTSDFWGGAGSACQGFITQLGRN 60

QY 54 ---VQCKWDATATLNNALQNLARTISEAGQAWA 84
DB 61 FQVIYEQANAHGQKVQAGNNAQTDASVGSWA 94

RESULT 25

H96904
uncharacterized small conserved protein, homolog of yfjA/yuke B. subtilis CAC0040 [impc
C:Species: Clostridium acetobutylicum
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000
C:Accession: D75391
R:White, O.; Eisen, J.A.; Heideberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: D75391
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-526 <WHI>
A:Cross-references: GB:AE001992; GB:AE000513; NID:96459233; PIDN:AAF11045.1; PID:9645923
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1480
A:Map position: 1

Query Match 14.9%; Score 71.5; DB 2; Length 526;
Best Local Similarity 23.8%; Pred. No. 31;
Matches 20; Conservative 16; Mismatches 37; Indels 11; Gaps 1;
QY 13 AAASAIQGNVTSIHSLDEGKQSLTKLAAAGGSGSEAYQGVQ-----QKWDAT 61
DB 172 ALVSLQOAGEDLRDEVDQRQAVEKAATKARKSGNALLSGLQNRABDLDRAGRDVDT 231

QY 62 ATELNNALQNLARTISEAGQAWAS 85

DB 232 RKDARQALMSAQKTAARKGNALLS 255

RESULT 26

AG1739
transmembrane protein [imported] - Listeria innocua (strain Clip11262)
C:Species: Listeria innocua
C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001
C:Accession: AG1739
R:Glaser, E.; Frangeul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker
D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihi, H.
Science 294, 849-852, 2001
A:Authors: Kref, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madueno, E.; Maitournam, A.; Ma
ok, C.; Schluter, T.; Simoes, N.; Tierrez, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland,
A:Title: Comparative genomics of Listeria species.
A:Reference number: AB1077; MUID:21537279; PMID:11679669
A:Accession: AG1739
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-927 <GLA>
A:Cross-references: GB:AL592022; PIDN:CAC97687.1; PID:gl6414982; GSPDB:GN00178
A:Experimental source: strain Clip11262
C:Genetics:
A:Gene: lin2460

Query Match 14.9%; Score 71.5; DB 2; Length 927;
Best Local Similarity 29.1%; Pred. No. 60;
Matches 30; Conservative 11; Mismatches 43; Indels 19; Gaps 5;

QY 1 MTEQQWNPAGIAAASAIQGNVT----SIHSLD-EGKQSLTKLAAAGG-----SGSEA 50
DB 480 VTELQGTGVAKTSAGYKAVHGGTTEAFNGIHAINGSQTLL-----GGANQLTAGLKS 533

QY 51 YQGVQCKWDATATLNNALQNLARTISEAGQAWASTGNVTGM 93

DB 534 AQAGNAKLIVAGTINELNNQIPTLT----SGVQLAAGSSNMENCL 573

RESULT 27

H96904
uncharacterized small conserved protein, homolog of yfjA/yuke B. subtilis CAC0040 [impc
C:Species: Clostridium acetobutylicum
C:Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 14-Sep-2001
C:Accession: H96904
R:Nolling, J.; Breton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee
J.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A:Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Cl
A:Reference number: A36900; MUID:21359325; PMID:21359325
A:Accession: H96904
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-97 <KUR>
A:Cross-references: GB:AE001437; PIDN:AAK78027.1; PID:gl5022861; GSPDB:GN00168
A:Experimental source: Clostridium acetobutylicum ATCC824
C:Genetics:
A:Gene: CAC0040

Query Match 14.8%; Score 71; DB 2; Length 97;
Best Local Similarity 20.0%; Pred. No. 4.7;
Matches 16; Conservative 18; Mismatches 46; Indels 0; Gaps 0;
QY 11 IEAASAIQGNVTSIHSLDEGKQSLTKLAAAGGSGSEAYQGVQCKWDATATLNNALQ 70
DB 12 IDHTASVFAKNGQDIENVLKALTEKNKLVSSWEGDAKAFSSEDFQAPQVTOFAKLVE 71
QY 71 NLARTISEAGQAWASTEGNV 90
DB 72 QIGQOLKSASKTMDQTDKAV 91

RESULT 28

S02017
regulatory protein pila - Neisseria gonorrhoeae
C:Species: Neisseria gonorrhoeae
C:Date: 01-Dec-1989 #sequence_revision 01-Dec-1989 #text_change 02-Sep-2000
C:Accession: S02017
R:Taba, M.K.; So, M.; Seifert, H.S.; Billyard, E.; Marchal, C.
EMBO J. 7, 4367-4378, 1988
A:Title: Pilin expression in Neisseria gonorrhoeae is under both positive and negative
A:Reference number: S02017; MUID:89210824; PMID:2854063
A:Accession: S02017
A:Molecule type: DNA
A:Residues: 1-356 <TAH>
A:Cross-references: EMBL:X13965; NID:g44835; PIDN:CMA32144.1; PID:g44836
C:Genetics:
A:Gene: pila
C:Superfamily: cell division protein ftsY
F:29-48/Domain: DNA binding #status predicted <DNA>

Query Match 14.8%; Score 71; DB 2; Length 356;
Best Local Similarity 22.6%; Pred. No. 22;
Matches 19; Conservative 15; Mismatches 50; Indels 0; Gaps 0;

QY 9 AGIEAASAIQGNVTSIHSLDEGKQSLTKLAAAGGSGSEAYQGVQCKWDATATLNNAL 68
DB 19 AQVQETAAKVESEVAQIVGNIKEDVESLAESVKGRAESAVETVSGAVEQVETVAEMPSE 78
QY 69 LQNLARTISEAGQAWASTEGNV 92
DB 79 AGEAAERVESAKEAVAETVGEAVG 102

RESULT 29

B70930
hypothetical protein Rv1793 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:25:57 ; Search time 4.52381 Seconds
(without alignments)
987.561 Million cell updates/sec

Title: US-09-805-427A-1
Perfect score: 479
Sequence: 1 MTEQQWNPAGIEAAASAIQG.....ISEACQAMASTEGNVTGMFA 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127825

Minimum DB seq length: 6
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	474	99.0	94	1	ESA6_MYCTU
2	186	38.8	95	1	ESA6_MYCTU
3	96	20.0	96	1	ESA6_MYCTU
4	81	16.9	96	1	ESA6_MYCTU
5	80.5	16.8	90	1	ESA6_MYCTU
6	80	16.7	95	1	CPFP MYCTU
7	80	16.7	1037	1	FLIC ECOLI
8	75	15.7	1037	1	ACRD ECOLI
9	75	15.7	5327	1	MACF MOUSE
10	72	15.0	5430	1	MACF HUMAN
11	72	15.0	5938	1	MAC4 HUMAN
12	71.5	14.9	94	1	ESA6_MYCTU
13	71.5	14.9	139	1	CSMC CHLITE
14	71	14.8	416	1	FTSY NEIGO
15	70.5	14.7	94	1	ESA6_MYCTU
16	70.5	14.7	94	1	ESA6_MYCTU
17	70.5	14.7	94	1	ESA6_MYCTU
18	70.5	14.7	139	1	CSMC CHLVI
19	70.5	14.7	388	1	MRP4 STRPY
20	70.5	14.7	2073	1	FAS1 SCHPO
21	70	14.6	103	1	ESA6_MYCTU
22	70	14.6	649	1	VATI METMA
23	69.5	14.5	382	1	IPAC SHIDI
24	68	14.2	409	1	GUB_BACCI
25	68	14.2	486	1	CUG1 HUMAN
26	68	14.2	486	1	CUG1 MOUSE
27	68	14.2	706	1	ADDG HUMAN
28	68	14.2	2584	1	SPCQ HUMAN
29	67	14.0	550	1	FLIC SHIFL
30	67	14.0	1063	1	SPTS YEAST
31	67	14.0	1738	1	CO4 MOUSE
32	67	14.0	2349	1	TPR_HUMAN
33	67	14.0	3718	1	LMA5_MOUSE

34	66.5	13.9	227	1	DJB8_MOUSE
35	66.5	13.9	454	1	XYLX_PSEPU
36	66.5	13.9	457	1	Y162_UREPA
37	66	13.8	185	1	TD52_MOUSE
38	66	13.8	526	1	NABD_MYCLE
39	66	13.8	530	1	YA9A_SCHPO
40	65.5	13.7	137	1	MBH2_ECOLI
41	65.5	13.7	465	1	LIPP HUMAN
42	65.5	13.7	625	1	DNAK HALME
43	65.5	13.7	5171	1	BPEA_HUMAN
44	65	13.6	95	1	ES6X_MYCLE
45	65	13.6	307	1	LIPA_CHLEN
46	65	13.6	390	1	FLAA_BORER
47	65	13.6	430	1	AROA_MYCLE
48	65	13.6	692	1	MYG_PODCA
49	65	13.6	836	1	VG26_BPMD2
50	64.5	13.5	461	1	US45_LACLC
51	64.5	13.5	640	1	DNAK_VIBHA
52	64.5	13.5	890	1	MUTS_RICCN
53	64.5	13.5	918	1	YNUB_CABEL
54	64.5	13.5	1050	1	LKTA_ACTAC
55	64.5	13.5	3695	1	LMA5_HUMAN
56	64	13.4	189	1	CAYP_HUMAN
57	64	13.4	421	1	FTSY_NEIMC
58	64	13.4	712	1	PAL2_TORAC
59	64	13.4	1319	1	BCC1_ACERY
60	64	13.4	1392	1	CND1_MOUSE
61	63.5	13.3	228	1	DRP5_BPR69
62	63.5	13.3	328	1	SGS3_DROHR
63	63.5	13.3	483	1	COBQ_VIBPA
64	63.5	13.3	975	1	KINH_DROME
65	63.5	13.3	1816	1	LMA4_MOUSE
66	63	13.2	194	1	HAM1_CAUCR
67	63	13.2	348	1	CTH2_MOUSE
68	63	13.2	348	1	LPXD_NEIMB
69	63	13.2	390	1	CARP_SACFI
70	63	13.2	461	1	SELA_HAEIN
71	63	13.2	504	1	AMPX_VIBPR
72	63	13.2	705	1	STT3_HUMAN
73	63	13.2	705	1	STT3_MOUSE
74	63	13.2	778	1	HTR6_HALN1
75	63	13.2	778	1	HTR6_HALSA
76	63	13.2	6885	1	SNE2_HUMAN
77	62.5	13.0	312	1	CVSK_SYNY3
78	62.5	13.0	395	1	FLA2_RHIME
79	62.5	13.0	395	1	FLB2_RHIME
80	62.5	13.0	527	1	NCF2_BOVIN
81	62.5	13.0	635	1	DNAK_VIBCH
82	62.5	13.0	852	1	VTFR_HSVI1
83	62.5	13.0	879	1	ORP5_HUMAN
84	62.5	13.0	953	1	LKA3_PASHA
85	62	12.9	130	1	RR11_CVAPA
86	62	12.9	185	1	VATE_DEIRA
87	62	12.9	200	1	AKIP_MOUSE
88	62	12.9	242	1	DJB3_MOUSE
89	62	12.9	269	1	PANB_COREP
90	62	12.9	284	1	TPMT_CIOIN

ALIGNMENTS

RESULT 1
ESA6_MYCTU STANDARD; PRT; 94 AA.
ID _ESA6_MYCTU
AC Q57165; O84901;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE 6 kDa early secretory antigenic target (ESAT-6).
GN ESAT6 OR RV3875 OR MT3989 OR MT027.10.
OS Mycobacterium tuberculosis, and
OS Mycobacterium bovis.

Q9qv17 mus musculus
P23099 pseudomonas
Q9pvy2 ureaplasma
Q62393 mus musculus
Q49617 mycobacteri
Q07788 schizosacch
P07113 escherichia
P16233 homo sapien
Q9nhb9 halobacteri
Q94833 homo sapien
Q49946 mycobacteri
Q92774 chlamydia p
Q06064 bordetella
Q0cc13 mycobacteri
Q05000 podocoryne
P24220 mycobacteri
P22865 lactococcus
O87384 vibrio harv
Q92119 rickettsia
P34487 caenorhabdi
P14462 actinobacil
O15230 homo sapien
Q13938 homo sapien
O30391 neisseria m
P35513 nicotiana t
P19450 acetobacter
Q84224 mus musculus
O80164 bacterioph
P13730 drosophila
Q87hm1 vibrio para
P17210 drosophila
P97927 mus musculus
Q9abs4 caulobacter
Q9cyc5 mus musculus
P95377 neisseria m
P22929 saccharomyc
P43910 haemophilus
P22929 saccharomyc
Q01693 vibrio prot
P46977 homo sapien
P46978 mus musculus
Q9hr92 halobacteri
Q48319 halobacteri
Q8wxh0 homo sapien
P73410 synecocyst
P13118 rhizobium m
P13119 rhizobium m
O7775 bos taurus
Q34241 vibrio chol
Q00140 ictaluriid h
Q9h0x9 homo sapien
P55116 pasteurella
P48136 cyanophora
Q9rwh1 deinococcus
Q9dcj7 mus musculus
Q35723 mus musculus
Q8fua5 corynebacte
Q7068 ciona intes

CC -!- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL Outstation

CC at the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to license@isb-sib.ch).

CC

CC ENBL; U34848; AAC44033.1; --

CC ENBL; X79562; CAAS6099.1; --

CC ENBL; AL022120; CAAL1967.1; --

CC ENBL; AF420491; AAL16896.1; --

CC ENBL; AE007190; AAK48357.1; --

CC ENBL; AF004671; AAC83446.1; --

CC PIR; A70803; A70803.

CC TIGR; MT3989; --

CC TubercuList; RV3875; --

CC Antigen; Complete proteome.

CC INIT MET 0

CC SEQUENCE 94 AA; 9773 MW; 19245B0EC478BC84 CRC64;

CC

CC Query Match 99.0%; Score 474; DB 1; Length 94;

CC Best local Similarity 100.0%; Pred. No 9e-41;

CC Matches 94; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC

CC QY 2 TEQWNFAGIEAASAIQGNVTSIHSLLDEGKSLTKLAAAWGGSGSEAYGVQVKWDAT 61

CC DB 1 TEQWNFAGIEAASAIQGNVTSIHSLLDEGKSLTKLAAAWGGSGSEAYGVQVKWDAT 60

CC

CC QY 62 ATELNNALQNLARTISPAQOMASTEIGNVTGMFA 95

CC DB 61 ATELNNALQNLARTISPAQOMASTEIGNVTGMFA 94

CC

CC RESULT 2

CC ESA6 MYCLE STANDARD; PRT; 95 AA.

CC ID ESA6 MYCLE STANDARD; PRT; 95 AA.

CC AC Q50206; Q33083;

CC DT 30-MAY-2000 (Rel. 39, Created)

CC DT 30-MAY-2000 (Rel. 39, Last sequence update)

CC DT 16-CTT-2001 (Rel. 40, Last annotation update)

CC DE 6 kDa early secretory antigenic target homolog (ESAT-6-like protein

CC DE (L-ESAT).

CC GN ESAT6 OR ESX OR L45 OR ML0049 OR MLCB628.12C.

CC OS Mycobacterium leprae.

CC OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;

CC OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

CC CX NCBI_TaxID=1769;

CC RN [1]

CC RP SEQUENCE FROM N.A.

CC RA Wiles B., Notenboom T., Naafs B., Offringa R., Ottenhoff T.;

CC RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.

CC [2]

CC RP SEQUENCE FROM N.A.

CC RC STRAIN=TN.

CC RX MEDLINE=21128732; PubMed=11234002;

CC RA Cole S.T., Eiglmeier K., Parkhill J., James K.D., Thomson N.R.,

CC RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,

CC RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,

CC RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,

CC RA Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,

CC RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,

CC RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,

CC RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,

CC RA Barrell B.G.;

CC RT "Massive gene decay in the leprosy bacillus.";

CC RL Nature 409:1007-1011(2001).

CC -!- SUBCELLULAR LOCATION: Secreted (By similarity).

CC -!- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

corresponding to the 40.1-50.0 min region on the linkage map."; DNA Res. 3:379-392(1996).

[5] SEQUENCE OF 1-19 FROM N.A. MEDLINE=83238225; PubMed=6305924; Szekely E., Simon M.; "DNA sequence adjacent to flagellar genes and evolution of flagellar-phase variation."; J. Bacteriol. 155:74-81(1983).

[6] SEQUENCE OF 1-12. STRAIN=K12 / EMG2; MEDLINE=97443975; PubMed=9298646; Link A.J., Robison K., Church G.M.; "Comparing the predicted and observed properties of proteins encoded in the genome of *Escherichia coli* K-12."; Electrophoresis 18:1259-1313(1997).

[7] SEQUENCE OF 1-4. STRAIN=K12 / W3110; MEDLINE=98263247; PubMed=9600841; Wilkins M.R., Gasteiger E., Tonella L., Ou K., Tyler M., Sanchez J.-C., Gooley A.A., Walsh B.J., Bairoch A., Appel R.D., Williams K.L., Hochstrasser D.F.; "Protein identification with N and C-terminal sequence tags in proteome projects."; J. Mol. Biol. 278:599-608(1998)

CC -1- FUNCTION: FLAGELLIN IS THE SUBUNIT PROTEIN WHICH POLYMERIZES TO FORM THE FILAMENTS OF BACTERIAL FLAGELLA.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL FLAGELLIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; M4358; AAA23950.1; -
EMBL; X17440; CAA35488.1; -
EMBL; AB000285; AAC74990.1; -
EMBL; D90832; BAA15744.1; -
EMBL; D90833; BAA15751.1; -
EMBL; J01607; AAA92491.1; -
PIR; A37249; FLEC.
SWISS-2DPAGE; P04949; COLI.
EcoGene; EGI0321; flic.
InterPro; IPR001029; Flagellin C.
InterPro; IPR001492; Flagellin N.
Pfam; PF00700; Flagellin C; 1.
Pfam; PF00669; Flagellin N; 1.
PRINTS; PR00207; FLAGELLIN.
XW Flagella; Complete proteome.
INIT MET 0
CONFLICT 284 284 P -> L (IN REF. 2).
FT SEQUENCE 497 AA; 51163 MW; 24B5419C21C7B4E8 CRC64;
SQ

Query Match 16.7%; Score 80; DB 1; Length 95;
Best Local Similarity 30.4%; Pred. No. 1.4;
Matches 24; Conservative 9; Mismatches 38; Indels 8; Gaps 1;

QY 9 AGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAGSGSSSEAYQGVQOKWDATATELNNA 68
Db 393 AGEALTAVANGKTTDPLKALDDATASVDKFRSSLG-----AVQNRDLSAVTNLANT 444
QY 69 LQNLTARTISEAGQAMASTE 87
Db 445 TTNLSEASQRIQADYATE 463

RESULT 8
ACRD ECOLI

ACRD ECOLI STANDARD; PRT; 1037 AA.
 P2177; P76971; P77178; Q46715;
 01-MAR-1992 (Rel. 21, Created)
 01-NOV-1997 (Rel. 35, Last sequence update)
 16-OCT-2001 (Rel. 40, Last annotation update)
 Probable aminoglycoside efflux pump (Acriflavine resistance protein D).
 AC RD OR B2470.
 Escherichia coli.
 Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 Enterobacteriaceae; Escherichia.
 NCBI TaxID=562;
 (1) SEQUENCE FROM N.A.
 STRAIN=K12;
 Nilles M.L., Bertrand K.P.;
 Submitted (JUL-1994) to the EMBL/GenBank/DBJ databases.
 (2) SEQUENCE FROM N.A.
 Ma D., Cook D.N., Alberti M., Nikaido H., Hearst J.E.;
 Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
 (3) SEQUENCE FROM N.A.
 STRAIN=K12 / MG1655;
 MEDLINE=97426617; PubMed=9278503;
 Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
 Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
 Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
 Mau B., Shao Y.;
 "The complete genome sequence of Escherichia coli K-12.";
 Science 277:1453-1474(1997).
 (4) SEQUENCE FROM N.A.
 STRAIN=K12;
 MEDLINE=97349980; PubMed=9205837;
 Yamamoto Y., Aiba H., Baba T., Hayashi K., Inada T., Isono K.,
 Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,
 Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nashimoto H.,
 Oshima T., Oyama S., Saito N., Sampei G., Satoh Y., Sivasubram S.,
 Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,
 Yamagata S., Horiuchi T.;
 "Construction of a contiguous 874-Kb sequence of the Escherichia coli
 K-12 genome corresponding to 50.0-68.8 min on the linkage map and
 analysis of its sequence features.";
 DNA Res. 4:91-113(1997).
 (5) SEQUENCE OF 998-1037 FROM N.A.
 STRAIN=K12;
 MEDLINE=92355499; PubMed=1644752;
 Bouvier J., Richard C., Higgins W., Bogler O., Stragier S.;
 "Cloning, characterization, and expression of the dape gene of
 Escherichia coli.";
 J. Bacteriol. 174:5265-5271(1992).
 (6) FUNCTION.
 MEDLINE=20158894; PubMed=10692383;
 Rosenberg E.Y., Ma D., Nikaido H.;
 "Acrd of Escherichia coli is an aminoglycoside efflux pump.";
 J. Bacteriol. 182:1754-1756(2000).
 CC -!- FUNCTION: PARTICIPATES IN THE EFFLUX OF AMINOGLYCOSIDES. CONFERS
 RESISTANCE TO A VARIETY OF THESE SUBSTANCES.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
 (BY similarity).
 CC -!- SIMILARITY: BELONGS TO THE ACRB/ACRD/ACRF FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC

DR EMBL; U12598; AAA20584.1; -;
 DR EMBL; U10436; AAA74741.1; -;
 DR EMBL; AE000334; AAC75523.1; -;
 DR EMBL; D90875; BAA16344.1; -;
 DR EMBL; D90876; BAA16348.1; -;
 DR EMBL; X57403; CAA40663.1; -;
 DR PIR; B65022; B65022.
 DR EcoGene; EG10014; acrd.
 DR InterPro; IPR001036; Acrflvin_res.
 DR Pfam; PF00873; ACR_tran; 1.
 DR PRINTS; PRO0702; ACRIFLAVINRP.
 DR TIGRFAMs; TIGR00915; 2A0602; 1.
 KW Transmembrane; Inner membrane; Transport; Complete proteome.
 FT DOMAIN 1 9
 FT TRANSMEM 10 28
 FT DOMAIN 29 339
 FT TRANSMEM 340 359
 FT DOMAIN 360 365
 FT TRANSMEM 366 385
 FT DOMAIN 386 391
 FT TRANSMEM 392 413
 FT DOMAIN 414 441
 FT TRANSMEM 442 460
 FT DOMAIN 461 473
 FT TRANSMEM 474 496
 FT DOMAIN 497 537
 FT TRANSMEM 538 556
 FT DOMAIN 557 870
 FT TRANSMEM 871 890
 FT DOMAIN 891 896
 FT TRANSMEM 897 916
 FT DOMAIN 917 922
 FT TRANSMEM 923 944
 FT DOMAIN 945 971
 FT TRANSMEM 972 990
 FT DOMAIN 991 1003
 FT TRANSMEM 1004 1026
 FT DOMAIN 1027 1037
 FT CONFLICT 303 303
 FT CONFLICT 372 372
 FT CONFLICT 385 385
 FT CONFLICT 461 461
 FT CONFLICT 665 665
 FT CONFLICT 763 763
 FT CONFLICT 775 775
 FT CONFLICT 778 778
 SQ SEQUENCE 1037 AA; 113047 MW; 961611E1D24FD4E5 CRC64;
 Query Match 15.7%; Score 75; DB 1; Length 1037;
 Best Local Similarity 26.2%; Pred. No. 10;
 Matches 27; Conservative 18; Mismatches 38; Indels 20; Gaps 4;
 QY 7 NFAGIEA-----AASAIQGNVTSIHLL-----DEGQSILTKLAAWGGSGSEAYQG 53
 Db 48 NYPGASQTLNVTQVIEQNTGLDNLNMYMSQSSGTQASVT-LSFKAGTDPDEAVQ 106
 QY 54 VQCKWDATATLNNALONLARTISEAQ-----AMASTEQNV 90
 Db 107 VQNLQSMARKLQPAVQNGQVTRKGTGNTILTIATVSDGSM 149
 RESULT 9
 MACF_MOUSE STANDARD; PRT; 5327 AA.
 ID MACF_MOUSE
 AC G9QXZ0; P97394; P97395; P97396;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Microtubule-actin crosslinking factor 1 (Actin cross-linking family
 7).
 GN MACF1 OR MACF OR ACF7 OR ACFP7.
 OS Mus musculus (Mouse).

OC Fukuyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC STRAIN=BALB/c;
RX MEDLINE=20069791; PubMed=10601340;
RA Leung C.L., Sun D., Zheng M., Knowles D.R., Liem R.K.H.;
RT "Microtubule actin cross-linking factor (MACF): a hybrid of dystonin
RT and dystrophin that can interact with the actin and microtubule
RT cytoskeletons.";
RL J. Cell Biol. 147:1275-1286(1999).
RN [2]
RP PARTIAL SEQUENCE FROM N.A. (ISOFORMS 1, 2 AND 3).
RC STRAIN=BALB/c; TISSUE=Brain;
RX MEDLINE=97124842; PubMed=8954775;
RA Bernier G., Mathieu M., De Repentigny Y., Vidal S.M., Kothary R.;
RT "Cloning and characterization of mouse ACF7, a novel member of the
RT dystonin subfamily of actin binding proteins.";
RL Genomics 38:19-29(1996).
CC -!- FUNCTION: F-ACTIN-BINDING PROTEIN WHICH MAY PLAY A ROLE IN CROSS-
CC LINKING ACTIN TO OTHER CYTOSKELETAL PROTEINS. ALSO BINDS TO
CC MICROTUBULES.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=2;
CC IsoId=Q9QXZ0-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9QXZ0-2; Sequence=VSP_000717;
CC Note=Sequence incomplete;
CC Name=3;
CC IsoId=Q9QXZ0-3; Sequence=VSP_000719;
CC -!- TISSUE SPECIFICITY: EXPRESSED MAINLY IN LUNG, BRAIN, SPINAL CORD,
CC SKELETAL AND CARDIAC MUSCLE, AND SKIN.
CC -!- SIMILARITY: Contains 1 actin-binding domain.
CC -!- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Contains 1 SH3 domain.
CC -!- SIMILARITY: Contains 37 spectrin repeats.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF150755; AAD32244.1; -;
DR EMBL; U67203; AAC52988.1; -;
DR EMBL; U67204; AAC52989.1; -;
DR EMBL; U67205; AAC52990.1; -;
DR HSSP; Q01082; 1BKX.
DR MGD; MGI:108559; Macf1.
DR InterPro; IPR001589; Actbind actinin.
DR InterPro; IPR001715; Calponin-like.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003108; GAS2.
DR InterPro; IPR001452; SH3.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00307; CH; 2.
DR Pfam; PF00036; efhand; 2.
DR Pfam; PF02187; GAS2; 1.
DR Pfam; PF00435; spectrin; 27.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00033; CH; 2.
DR SMART; SM00054; Efb; 2.
DR SMART; SM00243; GAS2; 1.
DR SMART; SM00326; SH3; 1.
DR SMART; SM00150; SPEC; 36.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.

DR PROSITE; PS50021; CH; 2.
DR PROSITE; PS50002; SH3; FALSE_NEG.
DR PROSITE; PS00018; EF_HAND; 2.
KW Actin-binding; Cytoskeleton; Calcium-binding; Repeat; SH3 domain;
KW Alternative splicing.
FT DOMAIN 1 295 ACTIN-BINDING (BY SIMILARITY).
FT DOMAIN 78 181 CH 1.
FT REPEAT 194 295 CH 2.
FT REPEAT 314 355 SPECTRIN 1.
FT REPEAT 591 623 SPECTRIN 2.
FT REPEAT 680 784 SPECTRIN 3.
FT REPEAT 786 800 SPECTRIN 4.
FT DOMAIN 871 923 SH3.
FT REPEAT 1250 1272 SPECTRIN 5.
FT REPEAT 1287 1342 SPECTRIN 6.
FT REPEAT 1458 1534 SPECTRIN 7.
FT REPEAT 1593 1660 SPECTRIN 8.
FT REPEAT 1817 1886 SPECTRIN 9.
FT REPEAT 1934 2044 SPECTRIN 10.
FT REPEAT 2262 2282 SPECTRIN 11.
FT REPEAT 2376 2397 SPECTRIN 12.
FT REPEAT 2400 2509 SPECTRIN 13.
FT REPEAT 2557 2617 SPECTRIN 14.
FT REPEAT 2620 2727 SPECTRIN 15.
FT REPEAT 2730 2837 SPECTRIN 16.
FT REPEAT 2840 2944 SPECTRIN 17.
FT REPEAT 2946 2972 SPECTRIN 18.
FT REPEAT 2986 3029 SPECTRIN 19.
FT REPEAT 3086 3162 SPECTRIN 20.
FT REPEAT 3195 3273 SPECTRIN 21.
FT REPEAT 3276 3382 SPECTRIN 22.
FT REPEAT 3385 3491 SPECTRIN 23.
FT REPEAT 3494 3518 SPECTRIN 24.
FT REPEAT 3603 3709 SPECTRIN 25.
FT REPEAT 3722 3817 SPECTRIN 26.
FT REPEAT 3869 3930 SPECTRIN 27.
FT REPEAT 3933 4039 SPECTRIN 28.
FT REPEAT 4042 4149 SPECTRIN 29.
FT REPEAT 4152 4258 SPECTRIN 30.
FT REPEAT 4361 4368 SPECTRIN 31.
FT REPEAT 4371 4477 SPECTRIN 32.
FT REPEAT 4480 4587 SPECTRIN 33.
FT REPEAT 4594 4695 SPECTRIN 34.
FT REPEAT 4698 4804 SPECTRIN 35.
FT REPEAT 4807 4872 SPECTRIN 36.
FT REPEAT 4917 4941 SPECTRIN 37.
FT CA_BIND 4987 4998 EF-HAND 1 (POTENTIAL).
FT CA_BIND 5023 5034 EF-HAND 2 (POTENTIAL).
FT DOMAIN 5173 5180 POLY-SER.
FT VARSPLIC 5252 5267 4 X 4 AA TANDEM REPEATS OF [GS]-S-R-[AR].
FT 1 181 MSSDEETLSRSCRSRSCRSRSCRSRSGSLSPCPGPD
FT TLPNWLPHKQKRSQSVLDPAERAVRVVADDERDQVKK
FT TPTWKVNLKMKVKRHINDLYEDRDGNLNLLEVLGGIK
FT LPREKGRMFRHQNQVQIALDFLQKQVQLVNIREDITDG
FT NPXLTGLIWTILHFQ -> MGNSLGCVKEPKESIAVPEK
FT API-SPKRRVFRKRGKILTPASHREALEGTGVIET
FT ETLTKTARLPKPGVGAHPDIFLPGDSAPNSGVGQ
FT GMIVQKESFOARIOTAHILLENESSVVGASWLEGMV
FT IAHLLONPAERNEKSVQLVEPRTRSCSRVALLPLQGE
FT TAVEQGTLLRHRSSTLPTDYPSETVDQDPSEGVSIG
FT GRTKSVPSAPSGDLVSKVASSIPKQSGDPIHTEPHVG
FT LVSCGKPIMPASQDLVSGITVILPSSSGYSGDGLRLHG
FT IRPDETEPKTSTPFSEEDGTLSLE (in isoform 3).
FT /FTid=VSP_000718.
FT 393 393 L -> P (IN REF. 2; AAC52990).
FT 1882 1882 Q -> H (IN REF. 2).
FT 5327 5327 AA; 607972 MW; 5DE4FF5A6514BFFA CRC64;
FT 15.7%; Score 75; DB 1; Length 5327;
Query Match

TISSUE=Brain;
 MEDLINE=98116662; PubMed=9455484;
 Seki N., Ohira M., Nagase T., Ishikawa K.-I., Miyajima N.,
 Nakajima D., Nomura N., Ohara O.;
 "Characterization of cDNA clones in size-fractionated cDNA libraries
 from human brain."
 DNA Res. 4:345-349 (1997).
 CC -!- FUNCTION: F-ACTIN-BINDING PROTEIN WHICH MAY PLAY A ROLE IN CROSS-
 LINKING ACTIN TO OTHER CYTOSKELETAL PROTEINS. ALSO BINDS TO
 MICROTUBULES (BY SIMILARITY).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=4;
 CC Name=2;
 CC IsoId=Q9UPN3-2; Sequence=Displayed;
 CC Name=1;
 CC IsoId=Q9UPN3-1; Sequence=VSP_007341;
 CC Name=3;
 CC IsoId=Q9UPN3-3; Sequence=Not described;
 CC Name=4;
 CC IsoId=Q96PK2-1; Sequence=External;
 CC -!- TISSUE SPECIFICITY: Ubiquitously expressed.
 CC -!- SIMILARITY: Belongs to the plakin or cytolinker family.
 CC -!- SIMILARITY: Contains 1 actin-binding domain.
 CC -!- SIMILARITY: Contains 1 calponin-homology (CH) domains.
 CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
 CC -!- SIMILARITY: Contains 1 SH3 domain.
 CC -!- SIMILARITY: Contains 37 spectrin repeats.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on
 use by non-profit institutions as long as its content is in
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announcements>
 or send an email to license@isb-sib.ch).
 CC
 CC EMBL; AB029290; BA83821.1; -
 CC EMBL; AF141968; AAF06360.1; -
 CC EMBL; AF325341; AAL39000.1; -
 CC EMBL; AF325330; AAL39000.1; JOINED.
 CC EMBL; AF325331; AAL39000.1; JOINED.
 CC EMBL; AF325332; AAL39000.1; JOINED.
 CC EMBL; AF325333; AAL39000.1; JOINED.
 CC EMBL; AF325334; AAL39000.1; JOINED.
 CC EMBL; AF325335; AAL39000.1; JOINED.
 CC EMBL; AF325336; AAL39000.1; JOINED.
 CC EMBL; AF325339; AAL39000.1; JOINED.
 CC EMBL; AF325340; AAL39000.1; JOINED.
 CC EMBL; AB033077; BA86565.1; -
 CC EMBL; AL137853; CAC15920.1; -
 CC EMBL; AB007934; BAA32310.2; -
 CC PIR; T00079; T00079.
 CC HSPG; Q01082; 1BKR.
 CC Genew; HGNC:13664; MACF1.
 CC GO; GO:0005856; C:cytoskeleton; NAS.
 CC GO; GO:0003780; F:actin cross-linking activity; NAS.
 CC GO; GO:0005509; F:calcium ion binding activity; NAS.
 CC GO; GO:0008017; F:microtubule binding activity; NAS.
 CC InterPro; IPR001589; Actbind actnin.
 CC InterPro; IPR001715; Calponin-like.
 CC InterPro; IPR002048; EF-hand.
 CC InterPro; IPR003108; GAS2.
 CC InterPro; IPR001452; SH3.
 CC InterPro; IPR002017; Spectrin.
 CC Pfam; PF00307; CH; 2.
 CC Pfam; PF00036; ehand; 2.
 CC Pfam; PF02187; GAS2; 1.
 CC Pfam; PF00435; spectrin; 27.
 CC ProDom; PD000012; EF-hand; 1.
 CC SMART; SMO00033; CH; 2.
 CC SMART; SMO00054; EFh; 2.
 CC SMART; SMO0243; GAS2; 1.
 CC

DR SMART; SM00150; SPEC; 36.
DR PROSITE; PS00019; ACTININ_1; 1.
DR PROSITE; PS00020; ACTININ_2; FALSE_NEG.
DR PROSITE; PS00021; CH; 2.
DR PROSITE; PS00018; EF_HAND; 2.
DR PROSITE; PS00002; SH3; FALSE_NEG.
KW Actin-binding; Cytoskeleton; Calcium; Calcium-binding; Repeat;
KW SH3 domain; Coiled coil; Alternative splicing.
FT DOMAIN 1 295 ACTIN-BINDING.
FT DOMAIN 78 181 CH 1.
FT DOMAIN 194 295 CH 2.
FT DOMAIN 243 265 COILED COIL (POTENTIAL).
FT DOMAIN 477 529 COILED COIL (POTENTIAL).
FT DOMAIN 722 751 COILED COIL (POTENTIAL).
FT DOMAIN 816 843 COILED COIL (POTENTIAL).
FT DOMAIN 1013 1118 COILED COIL (POTENTIAL).
FT DOMAIN 1164 1191 COILED COIL (POTENTIAL).
FT DOMAIN 1399 1690 COILED COIL (POTENTIAL).
FT DOMAIN 1780 1843 COILED COIL (POTENTIAL).
FT DOMAIN 1975 2005 COILED COIL (POTENTIAL).
FT DOMAIN 2039 2312 COILED COIL (POTENTIAL).
FT DOMAIN 2385 2417 COILED COIL (POTENTIAL).
FT DOMAIN 2544 2695 COILED COIL (POTENTIAL).
FT DOMAIN 2760 2838 COILED COIL (POTENTIAL).
FT DOMAIN 2911 3001 COILED COIL (POTENTIAL).
FT DOMAIN 3130 3164 COILED COIL (POTENTIAL).
FT DOMAIN 3244 3277 COILED COIL (POTENTIAL).
FT DOMAIN 3418 3482 COILED COIL (POTENTIAL).
FT DOMAIN 3596 3666 COILED COIL (POTENTIAL).
FT DOMAIN 3786 3806 COILED COIL (POTENTIAL).
FT DOMAIN 3852 3931 COILED COIL (POTENTIAL).
FT DOMAIN 3967 3987 COILED COIL (POTENTIAL).
FT DOMAIN 4084 4218 COILED COIL (POTENTIAL).
FT DOMAIN 4343 4378 COILED COIL (POTENTIAL).
FT DOMAIN 4408 4437 COILED COIL (POTENTIAL).
FT DOMAIN 4468 4498 COILED COIL (POTENTIAL).
FT DOMAIN 4907 4935 COILED COIL (POTENTIAL).
FT DOMAIN 5044 5067 COILED COIL (POTENTIAL).
FT REPEAT 314 355 SPECTRIN 1.
FT REPEAT 591 623 SPECTRIN 2.
FT REPEAT 680 784 SPECTRIN 3.
FT REPEAT 786 800 SPECTRIN 4.
FT DOMAIN 871 923 SH3.
FT REPEAT 1250 1272 SPECTRIN 5.
FT REPEAT 1287 1342 SPECTRIN 6.
FT REPEAT 1455 1534 SPECTRIN 7.
FT REPEAT 1547 1659 SPECTRIN 8.
FT REPEAT 1815 1891 SPECTRIN 9.
FT REPEAT 1932 2042 SPECTRIN 10.
FT REPEAT 2260 2280 SPECTRIN 11.
FT REPEAT 2372 2395 SPECTRIN 12.
FT REPEAT 2398 2507 SPECTRIN 13.
FT REPEAT 2510 2618 SPECTRIN 14.
FT REPEAT 2621 2728 SPECTRIN 15.
FT REPEAT 2731 2838 SPECTRIN 16.
FT REPEAT 2841 2945 SPECTRIN 17.
FT REPEAT 2987 3024 SPECTRIN 18.
FT REPEAT 3136 3163 SPECTRIN 19.
FT REPEAT 3187 3274 SPECTRIN 20.
FT REPEAT 3277 3383 SPECTRIN 21.
FT REPEAT 3386 3492 SPECTRIN 22.
FT REPEAT 3495 3601 SPECTRIN 23.
FT REPEAT 3604 3673 SPECTRIN 24.
FT REPEAT 3713 3819 SPECTRIN 25.
FT REPEAT 3832 3927 SPECTRIN 26.
FT REPEAT 3982 4043 SPECTRIN 27.
FT REPEAT 4046 4152 SPECTRIN 28.
FT REPEAT 4155 4262 SPECTRIN 29.

Query Match 15.0%; Score 72; DB 1; Length 5430;
Best Local Similarity 27.3%; Pred. No. 1.4e+02;
Matches 24; Conservative 14; Mismatches 28; Indels 22; Gaps 3;

QY 16 SAIQNVTSIHLLDEKQSLTKLAAWGGSGSEAYQGV---QQKWDAT----- 61
DB 4422 SQLKAKEETYQLLDKGR-LWLLSRDSDSGSGKTEQSVALLQKHVVSSKWEERKSKL 4479
QY 62 -----ATELNALONLARTISEAQAM 83
DB 4480 BEALNLATEFQNSLQEFINWLTLAQSL 4507
RESULT 11
MAC4_HUMAN
ID MAC4_HUMAN STANDARD; PRT; 5938 AA.
AC Q96PK2; Q8WXY1;
DT 15-SEP-2003 (Rel. 42, Created)
DT 15-SEP-2003 (Rel. 42, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Microtubule-actin crosslinking factor 1, isoform 4.
GN MACF1 OR ACF7 OR ABP620 OR KIAA0465 OR KIAA1251.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.; ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
RX MEDLINE=21833812; PubMed=11845288;
RA Gong T.-W.L., Besirli C.G., Lomax M.I.;
RT Mamm. Genome 12:852-861(2001).
CC -!- FUNCTION: May play a role in cross-linking cytoskeletal proteins
by binding intermediate filaments to the N-terminal plectin
repeats and microtubules to the C-terminus.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=4;
CC Name=4;
CC IsoId=Q96PK2-1; Sequence=Displayed;
CC Name=1;
CC IsoId=Q9UPN3-1; Sequence=External;
CC Name=2;
CC IsoId=Q9UPN3-2; Sequence=External;
CC Name=3;
CC IsoId=Q9UPN3-3; Sequence=External;
CC -!- TISSUE SPECIFICITY: Expressed in heart, lung, pituitary and
placenta, not found in brain, kidney, liver, pancreas or skeletal
muscle.
CC -!- SIMILARITY: Belongs to the plakin or cytolinker family.
CC -!- SIMILARITY: Contains 2 EF-hand calcium-binding domains.
CC -!- SIMILARITY: Contains 19 plectrin repeats.
CC -!- SIMILARITY: Contains 32 spectrin repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF317696; AAL09459.1; -
DR EMBL; AF325341; AAL38997.1; -
DR EMBL; AF325333; AAL38997.1; JOINED.
DR EMBL; AF325334; AAL38997.1; JOINED.
DR EMBL; AF325335; AAL38997.1; JOINED.
DR EMBL; AF325336; AAL38997.1; JOINED.
DR EMBL; AF325339; AAL38997.1; JOINED.
DR EMBL; AF325340; AAL38997.1; JOINED.
DR GO; GO:0008017; F: microtubule binding; ISS.
DR InterPro; IPR002048; EF-hand.
DR InterPro; IPR003108; GAS2.
DR InterPro; IPR001101; Plectrin_repeat.
DR InterPro; IPR002017; Spectrin.
DR Pfam; PF00036; ehand; 2.
DR Pfam; PF02187; GAS2; 1.

DR Pfam; PF00681; Plectin; 11.
DR Pfam; PF00435; spectrin; 26.
DR ProDom; PD000012; EF-hand; 1.
DR SMART; SM00054; EFh; 2.
DR SMART; SM00243; GAS2; 1.
DR SMART; SM00250; PLEC; 19.
DR SMART; SM00150; SPEC; 32.
DR PROSITE; PS00018; EF_HAND; 2.
KW Cytoskeleton; Calcium; Calcium-binding; Repeat; Coiled coil;
KW Alternative splicing.
FT DOMAIN 1830 1936 COILED COIL (POTENTIAL).
FT DOMAIN 2001 2192 COILED COIL (POTENTIAL).
FT DOMAIN 2282 2345 COILED COIL (POTENTIAL).
FT DOMAIN 2477 2507 COILED COIL (POTENTIAL).
FT DOMAIN 2541 2654 COILED COIL (POTENTIAL).
FT DOMAIN 2686 2814 COILED COIL (POTENTIAL).
FT DOMAIN 2887 2919 COILED COIL (POTENTIAL).
FT DOMAIN 3046 3197 COILED COIL (POTENTIAL).
FT DOMAIN 3262 3503 COILED COIL (POTENTIAL).
FT DOMAIN 3632 3666 COILED COIL (POTENTIAL).
FT DOMAIN 3746 3779 COILED COIL (POTENTIAL).
FT DOMAIN 3920 3984 COILED COIL (POTENTIAL).
FT DOMAIN 4098 4168 COILED COIL (POTENTIAL).
FT DOMAIN 4288 4308 COILED COIL (POTENTIAL).
FT DOMAIN 4354 4386 COILED COIL (POTENTIAL).
FT DOMAIN 4397 4433 COILED COIL (POTENTIAL).
FT DOMAIN 4469 4489 COILED COIL (POTENTIAL).
FT DOMAIN 4586 4720 COILED COIL (POTENTIAL).
FT DOMAIN 4845 4880 COILED COIL (POTENTIAL).
FT DOMAIN 4910 4939 COILED COIL (POTENTIAL).
FT DOMAIN 4970 5000 COILED COIL (POTENTIAL).
FT DOMAIN 5409 5437 COILED COIL (POTENTIAL).
FT DOMAIN 5546 5569 COILED COIL (POTENTIAL).
FT CA BIND 5598 5610 EF-HAND 1 (POTENTIAL).
FT CA BIND 5634 5646 EF-HAND 2 (POTENTIAL).
FT REPEAT 12 49 PLECTIN 1.
FT REPEAT 53 88 PLECTIN 2.
FT REPEAT 89 126 PLECTIN 3.
FT REPEAT 130 164 PLECTIN 4.
FT REPEAT 166 202 PLECTIN 5.
FT REPEAT 203 240 PLECTIN 6.
FT REPEAT 243 278 PLECTIN 7.
FT REPEAT 279 316 PLECTIN 8.
FT REPEAT 318 354 PLECTIN 9.
FT REPEAT 375 762 PLECTIN 10.
FT REPEAT 763 800 PLECTIN 11.
FT REPEAT 801 838 PLECTIN 12.
FT REPEAT 839 876 PLECTIN 13.
FT REPEAT 897 934 PLECTIN 14.
FT REPEAT 935 972 PLECTIN 15.
FT REPEAT 975 1011 PLECTIN 16.
FT REPEAT 1012 1049 PLECTIN 17.
FT REPEAT 1121 1158 PLECTIN 18.
FT REPEAT 1159 1196 PLECTIN 19.
FT REPEAT 1925 2032 SPECTRIN 1.
FT REPEAT 2052 2160 SPECTRIN 2.
FT REPEAT 2211 2313 SPECTRIN 3.
FT REPEAT 2320 2430 SPECTRIN 4.
FT REPEAT 2437 2543 SPECTRIN 5.
FT REPEAT 2550 2652 SPECTRIN 6.
FT REPEAT 2659 2758 SPECTRIN 7.
FT REPEAT 2765 2896 SPECTRIN 8.
FT REPEAT 2903 3008 SPECTRIN 9.
FT REPEAT 3015 3119 SPECTRIN 10.
FT REPEAT 3126 3229 SPECTRIN 11.
FT REPEAT 3236 3339 SPECTRIN 12.
FT REPEAT 3346 3445 SPECTRIN 13.
FT REPEAT 3453 3555 SPECTRIN 14.
FT REPEAT 3562 3664 SPECTRIN 15.
FT REPEAT 3671 3775 SPECTRIN 16.
FT REPEAT 3782 3884 SPECTRIN 17.
FT REPEAT 3891 3993 SPECTRIN 18.
FT REPEAT 4000 4102 SPECTRIN 19.

FT REPEAT 4109 4211 SPECTRIN 20.
FT REPEAT 4218 4320 SPECTRIN 21.
FT REPEAT 4327 4428 SPECTRIN 22.
FT REPEAT 4438 4544 SPECTRIN 23.
FT REPEAT 4551 4653 SPECTRIN 24.
FT REPEAT 4660 4763 SPECTRIN 25.
FT REPEAT 4770 4872 SPECTRIN 26.
FT REPEAT 4879 4982 SPECTRIN 27.
FT REPEAT 4989 5091 SPECTRIN 28.
FT REPEAT 5098 5201 SPECTRIN 29.
FT REPEAT 5208 5309 SPECTRIN 30.
FT REPEAT 5316 5418 SPECTRIN 31.
FT REPEAT 5425 5555 SPECTRIN 32.
FT CONFLICT 1712 1712 P -> S (IN REF. 1; AAL38997).
SQ SEQUENCE 5938 AA; 670134 MW; B87841112752DA004 CRC64;

Query Match 15.0%; Score 72; DB 1; Length 5938;
Best Local Similarity 27.3%; Pred. No. 1.5e+02;
Matches 24; Conservative 14; Mismatches 28; Indels 22; Gaps 3;
QY 16 SAIQGNVTSIHSLDEGKSLTKLAAAGGSGSEAYQGV---QQKWDA----- 61
DB 4924 SOLKAEETYNQLLDKGR--LMLSRDSDSGSKTEQSVALLQKWHVSSKMERKSKL 4981
QY 62 -----ATELNALQNLTARTISEAQAM 83
DB 4982 EEALNLTAEFQNSLQEFINLTLAEQSL 5009

RESULT 12
ES64_MVCTU
ID ES64_MVCTU STANDARD; PRT; 94 AA.
AC 005300;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Putative ESAT-6 like protein 4.
GN RV1198 OR MT1236 OR MTC1364.10.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Ev;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
RA Horisby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sultston J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains,"
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE BSA76 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>;
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; X13965; CRA32144.1; ALT_FRAME.
 CC HSP; P10121; 1FTS.
 CC InterPro; IPR003593; AAA ATPase.
 CC InterPro; IPR004390; ABC_transpt_FtsY.
 CC InterPro; IPR000897; SRP54.
 CC Pfam; PF00448; SRP54; 1.
 CC Pfam; PF02881; SRP54 N; 1.
 CC ProDom; PD000819; SRP54; 1.
 CC SMART; SM00382; AAA; 1.
 CC TIGRFAMs; TIGR00064; ftsy; 1.
 CC PROSITE; PS00300; SRP54; 1.
 CC Signal recognition particle; GTP-binding; RNA-binding; Membrane;
 KW Cell division.
 FT NP_BIND 224 231 GTP (BY SIMILARITY).
 FT NP_BIND 304 308 GTP (BY SIMILARITY).
 FT NP_BIND 368 371 GTP (BY SIMILARITY).
 FT MUTAGEN 307 307 G->A: DECREASE IN GTPASE ACTIVITY; UNABLE
 FT TO COMPLEMENT THE LACK OF FTSY IN E.COLI;
 FT NO TRANSLOCATION ACTIVITY.
 SQ SEQUENCE 416 AA; 44359 MW; ABD977BAFF79E4ED CRC64;

Query Match 14.8%; Score 71; DB 1; Length 416;
 Best Local Similarity 22.6%; Pred. No. 8.8;
 Matches 19; Conservative 15; Mismatches 50; Indels 0; Gaps 0;
 QY 9 AGTAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGGSGSEAYQVQKWDATATLNNNA 68
 DB 19 AQVQETAAKVESEVAQVGNKEDVESLAEKSVGAESAVETVSGAVEQVKETVAEMPSE 78
 QY 69 LQNARTISEAQAMASTGNTG 92
 DB 79 AGEAARVESAKEVAETVGEAVG 102

RESULT 15
 ES61 MYCTU STANDARD; PRT; 94 AA.
 AC P96364; O08120; O08122; Q9L781;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Antigen Mb9.9B (ESAT-6 like protein 1).
 GN (RV1037C OR MT1066 OR MTCY10G2.12) AND
 GN (RV3619C OR MT3721 OR MTCY15C10.33 OR MTCY07H7B.03).
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Erdbmann;
 RX MEDLINE=20130298; PubMed=10662800;
 RA Alderson M.R., Benent T., Day C.H., Zhu L., Moles D., Skeiky Y.A.W.,
 RA Coler R., Lewinson D.M., Reed S.G., Dillon D.C.;
 RT "Expression cloning of an immunodominant family of Mycobacterium
 RT tuberculosis antigens using human CD4(+) T cells.";
 RL J. Exp. Med. 191:551-560(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / Oshkosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 RA Bishal W.;
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;
 CC or send an email to license@isb-sib.ch).

CC -----
 CC EMBL; AF226277; AAF32406.1; -;
 CC EMBL; Z92539; CAB06842.1; -;
 CC EMBL; Z95436; CAB08822.1; -;
 CC EMBL; AE006989; AAK45317.1; -;
 CC EMBL; AE007171; AAK48082.1; -;
 CC PIR; D70560; D70560.
 CC TIGR; MT1066; -;
 CC TIGR; MT3721; -;
 CC Tuberculist; RV1037C; -;
 CC Tuberculist; RV3619C; -;
 KW Hypothetical protein; Complete proteome.
 FT CONFLICT 20 20 Q -> L (IN REF. 1 AND 3; AAK48082).
 FT CONFLICT 23 23 S -> L (IN REF. 1 AND 3; AAK48082).
 SQ SEQUENCE 94 AA; 9833 MW; 8F971BB03789C57E CRC64;

Query Match 14.7%; Score 70.5; DB 1; Length 94;
 Best Local Similarity 23.7%; Pred. No. 1.8;
 Matches 22; Conservative 21; Mismatches 33; Indels 17; Gaps 3;
 QY 6 WNFAGIEAASAIQGNVTSIHSLDEGKQSLTKLAAA---WGSSEAYQV----- 53
 DB 5 YQFGVDVDAHGAMIRAAQAGSLEA---EHOAIISDLTASDFWGGAGSAAQCGFTTQIGRNF 61
 QY 54 ---VQKWDATATLNNALONLARTISEAQAVNA 84
 DB 62 QVIYEQANAHGQKQVQAAGNNMAQTDSAVGSSWA 94

RESULT 16
 ES65 MYCTU STANDARD; PRT; 94 AA.
 AC O53942;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Putative ESAT-6 like protein 5.
 GN RV1793 OR MT1842 OR MTV049.15.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the

```

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Taylor K.,
RA Sulston J.E., Taylor K., Whitehead S., Whitehead S., Whitehead S.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL022021; CA917714.1; -.
DR EMBL; AE007043; AAK46113.1; -.
DR PIR; B70930; B70930.
DR TIGR; MT1842; -.
DR Tuberculist; Rv1793; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 94 AA; 9942 MW; 1B7AB4913129C8D8 CRC64;

Query Match 14.7%; Score 70.5; DB 1; Length 94;
Best Local Similarity 25.5%; Pred. No. 1.8;
Matches 24; Conservative 20; Mismatches 31; Indels 19; Gaps 4;

QY 6 WNFAGIEAASAIQGNVTSIHSLLDEGKQSITK--LAAA--WGGSGSEA----- 50
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 5 YQFGVDVDAHGMIRQAAS-----LEAEHQAIVRDLVLAAGDFWGGAGSVACQEFITQLGRN 60
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 51 YQGVQKQWDATATLNALQNLAARTISEAGQAMA 84
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 61 FQVIYEQANHGQKQVQAAGNNMAQTDSAVGSSWA 94

RESULT 17
ID ES66 MYCTU STANDARD; PRT; 94 AA.
AC P95242;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Putative ESAT-6 like protein 6.
GN RV2346C OR MT2411 OR MTCY98.15C.
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RX MEDLINE=98295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekle A.F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares S., Taylor K.,
RA Sulston J.E., Taylor K., Whitehead S., Whitehead S., Whitehead S.,
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL022021; CA917714.1; -.
DR EMBL; AE007043; AAK46113.1; -.
DR PIR; B70930; B70930.
DR TIGR; MT1842; -.
DR Tuberculist; Rv1793; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 94 AA; 9942 MW; 1B7AB4913129C8D8 CRC64;

```

```

RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; Z83860; CAB06161.1; -.
DR EMBL; AE007081; AAK46704.1; -.
DR PIR; C70662; C70662.
DR TIGR; MT2411; -.
DR Tuberculist; Rv2346c; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 94 AA; 9954 MW; C740622D0BB7C8CA CRC64;

Query Match 14.7%; Score 70.5; DB 1; Length 94;
Best Local Similarity 25.5%; Pred. No. 1.8;
Matches 24; Conservative 20; Mismatches 31; Indels 19; Gaps 4;

QY 6 WNFAGIEAASAIQGNVTSIHSLLDEGKQSITK--LAAA--WGGSGSEA----- 50
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 5 YQFGVDVDAHGMIRQAQA-----GLLEAEHQAIVRDLVLAAGDFWGGAGSVACQEFITQLGRN 60
QY : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 51 YQGVQKQWDATATLNALQNLAARTISEAGQAMA 84
Dy : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 61 FQVIYEQANHGQKQVQAAGNNMAQTDSAVGSSWA 94

RESULT 18
ID CSMC_CHLVI STANDARD; PRT; 139 AA.
AC Q46466;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Chlorosome protein C (14 kDa chlorosome protein).
GN CSMC.
OS Chlorobium vibrioforme.
OC Bacteria; Chlorobi; Chlorobia; Chlorobiales; Chlorobiaceae;
OC Chlorobium.
OX NCBI_TaxID=1098;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=8327D;
RA Chung S., Frank G., Zuber H., Bryant D.A.;
RT "Genes encoding two chlorosome components from the green sulfur
RT bacteria Chlorobium vibrioforme strain 8327D and Chlorobium
RT tepidum."
RL Photosyn. Res. 41:261-275 (1994).
CC -!- SUBCELLULAR LOCATION: Chlorosome.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----
DR EMBL; U09867; AAA18795.1; -
KW Photosynthesis; Bacteriochlorophyll; Chlorosome.
SQ SEQUENCE 139 AA; 14102 MW; A7B6F1376A342C9E CRC64;

Query Match 14.7%; Score 70.5; DB 1; Length 139;
Best Local Similarity 19.4%; Pred. NO. 2.8;
Matches 21; Conservative 20; Mismatches 48; Indels 19; Gaps 2;

QY 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAGWGGSG-----47

DB 30 LTGSAIVGGLEAGSVVETVAGTVGSLVD-----ASGIGNMIGSGGVGTIDRVAIT 84

QY 48 -SEAYGVQVQKWDATATEINNALONLARTISEAGQAMASTEGNVTGMF 94

DB 85 VKDVSRSAGDLYNDVRNVENVGTGNAKAVGDAGSSASDAVNKIAGSF 132

RESULT 19

ID MRP4 STRPY STANDARD; PRT; 388 AA.
AC P30141;
DT 01-APR-1993 (Rel. 25, Created)
DT 01-APR-1993 (Rel. 25, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fibrinogen- and Ig-binding protein precursor (MRP protein).
GN MRP4.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AP4 / Serotype M4;
RX MEDLINE=92269579; PubMed=1598817;
RA Stenberg L., O'Toole P., Lindahl G.;
RT "Many group A streptococcal strains express two different
RT immunoglobulin-binding proteins, encoded by closely linked genes:
RT characterization of the proteins expressed by four strains of
RT different M-type".
RL Mol. Microbiol. 6:1185-1194 (1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AP4 / Serotype M4;
RX MEDLINE=92409576; PubMed=1528877;
RA O'Toole P., Stenberg L., Rissler M., Lindahl G.;
RT "Two major classes in the M protein family in group A streptococci."
RL Proc. Natl. Acad. Sci. U.S.A. 89:8661-8665 (1992).
CC -!- FUNCTION: BINDS IGG MOLECULES OF THE IGI, IG2 AND IG4 SUBCLASSES,
CC AND ALSO BINDS FIBRINOGEN.
CC -!- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (potential).
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; M87831; AAA26930.1; -
DR InterPro; IPR005877; Gpos_Y5IRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR006192; LPXTG.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; Y5IRK signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; Y5IRK signal; 1.
DR PROSITE; PS50847; GRAM_POS_ANCHORING; 1.

KW IgG-binding protein; Cell wall; Peptidoglycan-anchor; Repeat; Signal.
FT SIGNAL 1 41
FT CHAIN 42 357 FIBRINOGEN- AND IG-BINDING PROTEIN.
FT PROPEP 358 388 REMOVED BY SORTASE (POTENTIAL).
FT SITE 354 358 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD RES 357 357 AMIDE-LINKED TO CELL WALL (POTENTIAL).
SQ SEQUENCE 388 AA; 42244 MW; 3AB19E8B6CEDA722 CRC64;

Query Match 14.7%; Score 70.5; DB 1; Length 388;

Best Local Similarity 23.3%; Pred. No. 9.1;
Matches 20; Conservative 20; Mismatches 41; Indels 5; Gaps 1;

QY 1 MTEQOWNFAGIEAASAIQGNVTSIHSLDEKQSLTKLAAGWGGSGSEAVGVQVKWDA 60

DB 203 LTELQAKLDTAEAKAKLESQVTTLENLGSAKRELTDQAKLDAANAE-----KERLQS 257

QY 61 TATELNALONLARTISEAGQAMAST 86

DB 258 QAATLEKQLEATKKELADLOAKLAAT 283

RESULT 20

ID FAS1 SCHPO STANDARD; PRT; 2073 AA.
AC Q9UUG0; P78799;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Fatty acid synthase subunit beta [EC 2.3.1.86] [Includes: 3-
DE hydroxypalmitoyl-[acyl-carrier-protein] dehydratase (EC 4.2.1.61);
DE Enoyl-[acyl-carrier-protein] reductase [NADH] (EC 1.3.1.9); [Acyl-
DE carrier-protein] acetyltransferase (EC 2.3.1.38); [Acyl-carrier-
DE protein] malonyltransferase (EC 2.3.1.39); S-acyl fatty acid synthase
DE thioesterase [EC 3.1.2.14].
GN FAS1 OR SPAC926.09C.
OS Schizosaccharomyces pombe (Fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A., AND SUBUNIT.
RX MEDLINE=98359895; PubMed=9693066;
RA Niwa H., Katayama E., Yanagida M., Morikawa K.;
RT "Cloning of the fatty acid synthetase beta subunit from fission yeast,
RT coexpression with the alpha subunit, and purification of the intact
RT multifunctional enzyme complex".
RL Protein Expr. Purif. 13:403-413 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=972;
RX MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajadream M.A., Lyne M., Lyne R., Stewart A.,
RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mundall K., Murphy J., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
RA Rutherford J., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
RA Woodward J., Volckaert G., Aert R., Robben J., Grymonprez B.,
RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritz C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lehrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,

135 ALIGQTKFHGSALLDE---SSVKLAATFGGGYEDYFDELIELYEVYAPFAELIQLVLS 191

QY 71 NIARTISEAQQA 82
|:|:|

Db 192 KHLFTLSQNEQA 203

RESULT 21

ES6D_MYCTU STANDARD; PRF; 103 AA.

ID_ES6D_MYCTU AC 005440;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative ESAT-6 like protein 13.

GN RV3905C OR MT4024 OR MTCY15F10.06.

OS Mycobacterium tuberculosis.

OC Bacteria; Actinobacteridae; Actinobacteriales; Actinomycetales;

OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.

OX NCBI_TaxID=1773;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=H37RV;

RX MEDLINE=98295987; PubMed=9634230;

RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,

RA Gordon S.V., Eiglmier K., Gas S., Barry C.E. III, Tekalia F.,

RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,

RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,

RA Hornby S., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,

RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,

RA Rutter A., Seeger K., Skelton S., Squares S., Squares R.,

RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;

RT "Deciphering the biology of Mycobacterium tuberculosis from the

RT complete genome sequence.";

RL Nature 393:537-544 (1998).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=CDC 1551 / Oshkosh;

RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,

RA Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey R.,

RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,

RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,

RA Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and

RT laboratory strains.";

RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.

CC -!- SIMILARITY: BELONGS TO THE ESAT6 FAMILY.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC

CC EMBL; Z94121; CAB08096.1; --

DR EMBL; AR007193; AAK48388.1; --

DR EIR; B70600; B70600.

DR TIGR; MT4024; --

DR Tuberculist; RV3905C; --

DR Hypothetical protein; Complete proteome.

KW SEQUENCE 103 AA, 10460 MW; 3994E27A7EDDF02 CRC64;

SQ

Query Match 14.6%; Score 70; DB 1; Length 103;

Best Local Similarity 24.1%; Pred. No. 2.2;

Matches 19; Conservative 13; Mismatches 47; Indels 0; Gaps 0;

QY 9 AGIEAASAIQGNVTSTSIHSLIDEGKQSLTKLAAAGGSGSEAYQGVOQKWDATATELNN 68

Db 12 AVMOGFAASLDGAABHLAVQLAELDAQVQGMGLGWRGASGSYAWELWHRGAGEVQLG 71

QY 69 LONARTISEAGCAWASTE 87


```

RP SEQUENCE FROM N.A.
RC STRAIN=WL-12;
RX MEDLINE=90332423; PubMed=2377467;
RA Bueno A., Vasquez de Aldana C.R., Correa J., del Rey F.;
RT "Nucleotide sequence of a 1,3-1,4-beta-glucanase-encoding gene in
RL Bacillus circulans WL-12.";
CC Nucleic Acids Res. 18:4248-4248(1990).
CC -!- CATALYTIC ACTIVITY: Hydrolysis of 1,4-beta-D-glycosidic linkages
CC in beta-D-glucans containing 1,3- and 1,4-bonds.
CC -!- SIMILARITY: BELONGS TO CELLULOSE FAMILY D (FAMILY 8 OF GLYCOSYL
CC HYDROLASES).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X52880; CAA37062.1; -
CC PIR; S10485; S10485.
CC HSSP; P04955; ICSE.
CC InterPro; IPR002037; Glyco hydro 8.
CC Pfam; PF01270; Glyco hydro 8; 1.
CC PRINTS; PR00735; GLHYDRLASE8.
CC PROSITE; PS00812; GLYCOSYL_HYDROL_F8; 1.
CC Hydrolase; Glycosidase; Signal.
CC SIGNAL 1 31
CC CHAIN 32 409
CC BETA-GLUCANASE.
CC ACT SITE 95 95
CC PROTON DONOR (BY SIMILARITY).
CC ACT SITE 156 156
CC NUCLEOPHILE (POTENTIAL).
CC SEQUENCE 409 AA; 44816 MW; C75612167D13E76 CRC64;
CC
CC Query Match 14.2%; Score 68; DB 1; Length 409;
CC Best Local Similarity 38.8%; Pred. No. 17;
CC Matches 19; Conservative 5; Mismatches 17; Indels 8; Gaps 1;
CC
CC QY 3 EQWNFAQIEAAASAIQGNVTSIHSLLDEGKQSLTKLAAWGGSGSEAY 51
CC Db 140 KNSFSQNIQEGADSGDMDIAYSLLDKQ-----WGSGGSINY 180
CC
CC RESULT 25
CC CUG1_HUMAN
CC ID CUG1_HUMAN STANDARD; PRT; 486 AA.
CC AC Q92879; Q9NP83; Q9NR06;
CC DT 16-OCT-2001 (Rel. 40, Created)
CC DT 16-OCT-2001 (Rel. 40, Last sequence update)
CC DT 15-SEP-2003 (Rel. 42, Last annotation update)
CC DE CUG triplet repeat RNA-binding protein 1 (CUG-BP1) (RNA-binding
CC DE protein BRUNOL-2) (deadenylation factor CUG-BP) (50 kDa Nuclear
CC DE polyadenylated RNA-binding protein) (EDEN-BP).
CC GN CUGBP1 OR BRUNOL2 OR CUGBP OR NAB50.
CC OS Homo sapiens (Human).
CC OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
CC OX NCBI_TaxID=9606;
CC RN [1]
CC RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
CC RX MEDLINE=97105883; PubMed=8948631;
CC RA Timchenko L.T., Miller J.W., Timchenko N.A., DeVore D.R., Datar K.V.,
CC RA Lin L., Roberts R., Caskey C.T., Swanson M.S.;
CC RT "Identification of a (CUG)n triplet repeat RNA-binding protein and its
CC RT expression in myotonic dystrophy.";
CC RL Nucleic Acids Res. 24:4407-4414(1996).
CC RN [2]
CC RP SEQUENCE FROM N.A. (ISOFORM 1).
CC RX MEDLINE=20435814; PubMed=10893231;
CC RA Good P.J., Chen Q., Warner S.J., Herring D.C.;
CC RA "A family of human RNA-binding proteins related to the Drosophila
CC RT Bruno translational regulator.";
CC RL J. Biol. Chem. 275:28583-28592(2000).

```

```

RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 3).
RP TISSUE=Brain, and Skeletal muscle;
RC Takahashi N., Sasagawa N., Usuki F., Kawahara H., Sorimachi H.,
RA Maeda T., Suzuki K., Ishiura S.;
RT "The CUG-binding protein exists in multiple isoforms and reduces DM
RL protein kinase expression.";
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
CC SEQUENCE FROM N.A. (ISOFORM 1).
CC RA Paillard L., Legagneux V., Osborne B.;
CC RT "EDEN-BP/CUG-BP is a highly conserved deadenylation factor involved in
CC the post-translational regulation of c-jun proto-oncogene.";
CC RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC RN [5]
CC CHARACTERIZATION.
CC RX MEDLINE=20007898; PubMed=10536163;
CC RA Timchenko N.A., Welm A.L., Lu X., Timchenko L.T.;
CC RT "CUG repeat binding protein (CUGBP1) interacts with the 5' region of
CC C/EBPbeta mRNA and regulates translation of C/EBPbeta isoforms.";
CC RL Nucleic Acids Res. 27:4517-4525(1999).
CC RN [6]
CC CHARACTERIZATION.
CC RX MEDLINE=21269330; PubMed=11124939;
CC RA Timchenko N.A., Cai Z.J., Welm A.L., Reddy S., Ashizawa T.,
CC RT "RNA CUG repeats sequester CUGBP1 and alter protein levels and
CC activity of CUGBP1.";
CC RL J. Biol. Chem. 276:7820-7826(2001).
CC -!- FUNCTION: Regulates splicing and translation of various RNAs.
CC Binds to (CUG)n triplet repeats and to Bruno response elements.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS:
CC Event=Alternative splicing; Named isoforms=3;
CC Name=1; Synonyms=Lx1Q;
CC IsoId=Q92879-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q92879-2; Sequence=VSP_005784;
CC Name=3; Synonyms=A;
CC IsoId=Q92879-3; Sequence=VSP_005784, VSP_005785;
CC -!- TISSUE SPECIFICITY: Ubiquitous.
CC -!- DISEASE: CUGBP1 may be involved in the mechanism of myotonic
CC dystrophy. It binds to the CUG repeat expansion of the CC
CC untranslated region of the myotonin protein kinase (Mt-PK) gene.
CC -!- SIMILARITY: Contains 3 RNA recognition motif (RRM) domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U63289; AAC50895.1; -
CC EMBL; AF248648; AAF86230.1; -
CC EMBL; AF267533; AAF78955.1; -
CC EMBL; AF267534; AAF78956.1; -
CC EMBL; A007988; CAC20566.1; -
CC HSSP; P09651; 1HAI.
CC Genes; HGNC:2549; CUGBP1.
CC GK; Q92879; -
CC MIM; 601074; -
CC GO; GO:0005634; C:nucleus; TAS.
CC GO; GO:0003723; F:RNA binding activity; TAS.
CC GO; GO:0006397; P:mRNA processing; TAS.
CC InterPro; IPR000504; RNA_rec_mot.
CC Pfam; PF00076; rrm; 3.
CC SMART; SM00360; RRM; 3.
CC PROSITE; PS00102; RRM; 3.
CC PROSITE; PS00030; RRM_RNP_1; FALSE NEG.
CC mRNA processing; RNA-binding; Repeat; Nuclear protein;
CC Alternative splicing.

```

FT	DOMAIN	16	99	RNA-BINDING (RRM) 1.
FT	DOMAIN	108	188	RNA-BINDING (RRM) 2.
FT	DOMAIN	287	308	SER-RICH.
FT	DOMAIN	401	479	RNA-BINDING (RRM) 3.
FT	VARSPPLIC	231	234	Missing (in isoform 2 and isoform 3).
FT	VARSPPLIC	297	297	FTId=Vsp.005784.
FT	VARSPPLIC	297	297	S -> SA (in isoform 3).
FT	VARSPPLIC	297	297	FTId=VSP.005785.
SQ	SEQUENCE	486 AA;	52063 MW;	C4C13D77273A01D CRC64;
Query Match				
Best Local Similarity 23.8%; Pred. No. 21;				
Matches 34; Conservative 9; Mismatches 40; Indels 60; Gaps 4;				
Qy	4	QOWNPAGIEAAASATQGNVTSHSLDEGKQ	34	
Db	260	QLQNLAALAAASAAQNTPSGNTALTTSSPLSVLTSSGSPSSSSNVNPIASLGALQ	319	
Qy	35	-----SLTKLAAMGGSGS-----EAYGVQCKWDATATLNN	67	
Db	320	TLAGATAGLNVGSLAGMAALNGGLSGLSNGTGTMTGATQAYSGIQYAAAAALPTLYN	379	
Qy	68	ALQNL--ARTISEAGQAMASTEg 88		
Db	380	--QNLTOOSIGAAGSQXEGPEG 400		
RESULT 26				
ID	CUG1 MOUSE	STANDARD;	PRT;	486 AA.
AC	P28659; Q9CXES; Q9ERJ8; Q9J137;			
DT	01-DEC-1992 (Rel. 24, Created)			
DT	15-OCT-2001 (Rel. 40, Last sequence update)			
DE	CUG triplet repeat RNA-binding protein 1 (CUG-BP1) (RNA-binding			
DE	protein BRUNO-2) (Deadenylation factor CUG-BP) (Deadenylation factor			
DE	EDEN-BP) (Brain protein F41).			
GN	CUGBP1 OR BRUNO12 OR CUGBP.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxId=10090;			
RN	[1]_SEQUENCE FROM N.A. (ISOFORM 1).			
RP	STRAIN=Swiss; TISSUE=Ovary;			
RA	Faillard L., Legagneux V., Osborne B.;			
RT	"EDEN-BP/CUG-BP is a highly conserved deadenylation factor involved in			
RT	the post-translational regulation of c-jun proto-oncogene."			
RL	Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.			
RN	(2)			
RP	SEQUENCE FROM N.A. (ISOFORM 1).			
RC	TISSUE=Liver;			
RA	Takahashi N., Sasagawa N., Usuki F., Kawahara H., Sorimachi H.,			
RA	Maeda T., Suzuki K., Ishiura S.;			
RT	"The CUG-binding protein exists in multiple isoforms and reduces DM			
RT	protein kinase expression."			
RL	Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.			
RN	[3]			
RP	SEQUENCE FROM N.A. (ISOFORM 3).			
RC	STRAIN=C57BL/6J; TISSUE=Embryonic liver;			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,			
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,			
RA	Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,			
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Baren G.,			
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			

RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,			
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,			
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-P.,			
RA	Suzuki H., Iyoy-oka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,			
RA	Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,			
RA	Hayashizaki Y.;			
RT	"Functional annotation of a full-length mouse cDNA collection."			
RL	Nature 409:685-690(2001).			
RN	[4]			
RP	SEQUENCE OF 127-445 FROM N.A. (ISOFORM 1).			
RA	Kato K.;			
RL	Submitted (AUG-1991) to the EMBL/GenBank/DBJ databases.			
CC	!- FUNCTION: Regulates splicing and translation of various RNAs.			
CC	Binds to (CUG)n triplet repeats and to Bruno response elements (By			
CC	similarity).			
CC	!- SUBCELLULAR LOCATION: Nuclear (By similarity).			
CC	!- ALTERNATIVE PRODUCTS:			
CC	Event=Alternative splicing; Named isoforms=3;			
CC	Comment=Experimental confirmation may be lacking for some			
CC	isoforms;			
CC	Name=1; Synonyms=LVIQ;			
CC	Isoid=P28659-1; Sequence=Displayed;			
CC	Name=2;			
CC	Isoid=P28659-2; Sequence=VSP_005786;			
CC	Name=3; Synonyms=A;			
CC	Isoid=P28659-3; Sequence=VSP_005787;			
CC	!- TISSUE SPECIFICITY: NEOCORTEX, CEREBELLAR CORTEX, HIPPOCAMPUS AND			
CC	OTHER AREAS, ABUNDANT IN THE PUTAMEN, AND POORLY EXPRESSED IN THE			
CC	THALAMUS AND IN THE BRAIN STEM.			
CC	!- SIMILARITY: Contains 3 RNA recognition motif (RRM) domains.			
CC	!- CAUTION: Ref.4 sequence differs from that shown due to a			
CC	frameshift in position 367.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AJ007987; CAC20707.1; -			
DR	EMBL; AF267535; AAF78957.1; -			
DR	EMBL; AK014492; BAB29392.1; -			
DR	EMBL; X61451; CAA43691.1; ALT_FRAME.			
DR	MCD; MGI:1342295; Cugbp1.			
DR	GO; GO:0008248; P:pre-mRNA splicing factor activity; IDA.			
DR	GO; GO:0006376; P:RNA splice site selection; IDA.			
DR	InterPro; IPR000504; RNA_rec_mot.			
DR	Pfam; PF00076; rrm; 3.			
DR	PROSITE; PS0102; RRM; 3.			
DR	PROSITE; PS00030; RRM_RNP_1; FALSE NEG.			
KW	mRNA processing; RNA-binding; Repeat; Nuclear protein;			
KW	Alternative splicing.			
FT	DOMAIN	16	99	RNA-BINDING (RRM) 1.
FT	DOMAIN	108	188	RNA-BINDING (RRM) 2.
FT	DOMAIN	287	308	SER-RICH.
FT	DOMAIN	401	479	RNA-BINDING (RRM) 3.
FT	VARSPPLIC	231	234	Missing (in isoform 2 and isoform 3).
FT	VARSPPLIC	297	297	Missing (in isoform 2 and isoform 3).
FT	VARSPPLIC	297	297	S -> SA (in isoform 3).
FT	VARSPPLIC	297	297	FTId=VSP_005787.
FT	CONFLICT	291	291	L -> P (IN REF. 2).
FT	CONFLICT	301	301	P -> T (IN REF. 2).
FT	CONFLICT	335	335	G -> R (IN REF. 2).
FT	CONFLICT	347	347	G -> A (IN REF. 2).
SQ	SEQUENCE	486 AA;	52107 MW;	ABE22331A62B584 CRC64;
Query Match				
Best Local Similarity 14.2%; Score 68; DB 1; Length 486;				
Matches 34; Conservative 9; Mismatches 40; Indels 60; Gaps 4;				
Qy	4	QOWNPAGIEAAASATQGNVTSHSLDEGKQ	34	

DR	EMBL; Y14375;	CAB51805.1;	JOINED.
DR	EMBL; Y14376;	CAB51805.1;	JOINED.
DR	EMBL; Y14377;	CAB51805.1;	JOINED.
DR	EMBL; Y14378;	CAB51805.1;	JOINED.
DR	EMBL; Y14379;	CAB51805.1;	JOINED.
DR	EMBL; Y14380;	CAB51805.1;	JOINED.
DR	EMBL; Y14381;	CAB51805.1;	JOINED.
DR	EMBL; Y14382;	CAB51805.1;	JOINED.
DR	EMBL; Y14383;	CAB51805.1;	JOINED.
DR	EMBL; Y14384;	CAB51805.1;	JOINED.
DR	EMBL; Y14372;	CAB51806.1;	-
DR	EMBL; Y14373;	CAB51806.1;	JOINED.
DR	EMBL; Y14374;	CAB51806.1;	JOINED.
DR	EMBL; Y14375;	CAB51806.1;	JOINED.
DR	EMBL; Y14376;	CAB51806.1;	JOINED.
DR	EMBL; Y14377;	CAB51806.1;	JOINED.
DR	EMBL; Y14378;	CAB51806.1;	JOINED.
DR	EMBL; Y14379;	CAB51806.1;	JOINED.
DR	EMBL; Y14380;	CAB51806.1;	JOINED.
DR	EMBL; Y14381;	CAB51806.1;	JOINED.
DR	EMBL; Y14382;	CAB51806.1;	JOINED.
DR	EMBL; Y14383;	CAB51806.1;	JOINED.
DR	EMBL; Y14384;	CAB51806.1;	JOINED.
DR	GeneW; HGNC:245;	ADD3.	
DR	MIM; 601568;	--	
DR	GO; GO:0016020;	C-membrane; TAS.	
DR	GO; GO:0005200;	F-structural constituent of cytoskeleton; TAS.	
DR	InterPro; IPR001303;	Aldolase II N.	
DR	Pfam; PF00596;	Aldolase II; 1_--N.	
KW	Cytoskeleton; Membrane; Calmodulin-binding; Phosphorylation;		
KW	Alternative splicing.		
FT	DOMAIN	684 701	
FT	MOD RES	683 683	
FT	VARSPLIC	576 607	
FT	CONFLICT	49 50	CALMODULIN-BINDING (POTENTIAL).
FT	CONFLICT	60 60	PHOSPHORYLATION (BY PKC) (BY SIMILARITY).
FT	CONFLICT	362 362	Missing (in isoform 1).
FT	CONFLICT	420 420	/FTID-VSP 000188.
FT	CONFLICT	421 421	FN -> SS (IN REF. 1).
FT	CONFLICT	421 421	Q -> R (IN REF. 1).
FT	CONFLICT	421 421	Q -> P (IN REF. 2).
FT	CONFLICT	421 421	V -> M (IN REF. 2).
FT	CONFLICT	426 433	P -> L (IN REF. 1).
FT	CONFLICT	484 484	KYMAQRQQ -> QIHGTRGNK (IN REF. 2).
FT	CONFLICT	484 484	X -> Q (IN REF. 2).
SQ	SEQUENCE	706 AA;	EBBEAF602A4D7B41 CRC64;
	Query Match	14.2%;	Score 68; DB 1; Length 706;
	Best Local Similarity	31.4%;	Pred. No. 32;
	Matches 22;	Conservative 8;	Mismatches 34; Indels 6; Gaps 3;
QY	11 IEAA5AIGNTVTSIHLDEGQSLLTALAAGWSGSSEAYGGVOOKWDATATELNALQ 70	:	:
Dd	319 IQVQALAGAGGVNDLHFLDFQKYKAPTVAASGGGG--VNMGSHQKWVGIEI-----FE 373	:	:
QY	71 NLARTISEAG 80	:	:
Dd	373 GLMRTLDNLG 382	:	:
RESULT 28			
SPQC HUMAN			
ID	SPCO_HUMAN	STANDARD;	PRT; 2564 AA.
AC	Q9H254; Q9HIK7; Q9HIK8; Q9HIK9; Q9H3G8; Q9HCD0;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	Spectrin beta chain, brain 3 (Spectrin, non-erythroid beta chain 3)		
DE	(Beta-IV spectrin).		
GN	SPTEN4 OR SPTEN3 OR KIAA1642.		
OC	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
RN	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A. (ISOFORM 1).		
RP	MEDLINE=21316449; PubMed=11294830;		

RA Tse W.T., Tang J., Jin O., Korsgren C., John K.M., Kung A.L.,
RA Gwynn B., Peters L.B., Lux S.B.;
RT "A new spectrin, beta-IV, has a major truncated isoform that
RT associates with promyelocytic leukemia protein nuclear bodies and the
RT nuclear matrix.";
RL J. Biol. Chem. 276:23974-23985(2001).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORMS 1; 2; 3 AND 4).
RX MEDLINE=20539976; PubMed=11086001;
RA Berghs S., Aggujaro D., Dirx R. Jr., Maksimova E., Stabach P.,
RA Hermel J.-M., Zhang J.-P., Philbrick W., Slepnev V., Ort T.,
RA Solimena M.;
RT "Betalv spectrin, a new spectrin localized at axon initial segments
RT and nodes of ranvier in the central and peripheral nervous system.";
RL J. Cell Biol. 151:985-1002(2000).
RN [3]
RP SEQUENCE OF 386-2382 FROM N.A. (ISOFORM 1).
RX TISSUE=Brain;
RA MEDLINE=20450683; PubMed=10997877;
RA Nagase T., Kikuno R., Nakayama M., Hirose M., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes.
RT XVIII. The complete sequences of 100 new cDNA clones from brain which
RT code for large proteins in vitro.";
RL DNA Res. 7:273-281(2000).
CC -!- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=4;
CC Name=1;
CC IsoId=Q9H254-1; Sequence=Displayed;
CC Name=2;
CC IsoId=Q9H254-2; Sequence=VSP_000725, VSP_000726;
CC Name=3;
CC IsoId=Q9H254-3; Sequence=VSP_000723, VSP_000724;
CC Name=4;
CC IsoId=Q9H254-4; Sequence=VSP_000727, VSP_000728;
CC -!- TISSUE SPECIFICITY: Abundantly expressed in brain and pancreatic
CC islets.
CC -!- SIMILARITY: BELONGS TO THE SPECTRIN FAMILY.
CC -!- SIMILARITY: Contains 2 calponin-homology (CH) domains.
CC -!- SIMILARITY: Contains 1 PH domain.
CC -!- SIMILARITY: Contains 18 spectrin repeats.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF311855; AAG42473.1; -;
CC EMBL; AF082075; AAG38974.1; -;
CC EMBL; AY004226; AAF93171.1; -;
CC EMBL; AY004226; AAF93172.1; -;
CC EMBL; AY004227; AAF93173.1; -;
CC EMBL; AB046862; BAB13468.1; -;
CC HSP; Q01082; IAKR;
CC Genbank; HGNC:14896; SPTEN4.
CC MIM; 606214; -;
CC GO; GO:0016363; C:nuclear matrix; IDA.
CC GO; GO:0016605; C:PML body; IDA.
CC GO; GO:0008091; C:spectrin; IDA.
CC GO; GO:0003779; F:actin binding activity; TAS.
CC GO; GO:0005006; F:ankyrin binding activity; IDA.
CC GO; GO:0005200; F:structural constituent of cytoskeleton; TAS.
CC GO; GO:0007016; P:cytoskeletal anchoring; TAS.
CC GO; GO:0016192; P:vesicle-mediated transport; TAS.
CC InterPro; IPR001589; Actbind actnin.
CC InterPro; IPR001715; Calponin-like.
CC InterPro; IPR001849; PH.
CC InterPro; IPR002017; Spectrin.
CC InterPro; IPR001605; Spectrin_PH.
CC Pfam; PF00307; CH; 2.
CC Pfam; PF00169; PH; 1.

DR Pfam; PF00435; spectrin; 18.
DR PRINTS; PR00683; SPECTRINPH.
DR SMART; SM00033; CH; 2.
DR SMART; SMC0233; PH; 1.
DR SMART; SMC0150; SPEC; 16.
DR DR PROSITE; PS00019; ACTININ_1; 1.
DR DR PROSITE; PS00020; ACTININ_2; 1.
DR DR PROSITE; PS00021; CH; 2.
DR DR PROSITE; PS00003; PH DOMAIN; 1.
KW Cytoskeleton; Membrane; Repeat; Actin-binding; Capping protein;
KW Alternative splicing.
FT DOMAIN 1 282 ACTIN-BINDING (BY SIMILARITY).
FT CH 1.
FT CH 2.
FT SPECTRIN 1.
FT SPECTRIN 2.
FT SPECTRIN 3.
FT SPECTRIN 4.
FT SPECTRIN 5.
FT SPECTRIN 6.
FT SPECTRIN 7.
FT SPECTRIN 8.
FT SPECTRIN 9.
FT SPECTRIN 10.
FT SPECTRIN 11.
FT SPECTRIN 12.
FT SPECTRIN 13.
FT SPECTRIN 14.
FT SPECTRIN 15.
FT SPECTRIN 16.
FT SPECTRIN 17.
FT SPECTRIN 18.
FT PH.
FT Missing (in isoform 3).
FT /FTID=VSP_000723.
FT AVAAEGLRQNIYGEQAEAVTRLEK -> MHHYVSCS
FT SAPSGTPIPPQIQLEAHRH (in isoform 3).
FT /FTID=VSP_000724.
FT NOENLRQAQQMQLHDLQLQH -> CLIIHPALHPWE
FT PVLPRSSS (in isoform 2).
FT /FTID=VSP_000725.
FT Missing (in isoform 2).
FT /FTID=VSP_000726.
FT IBKIAEQSKQPTPLLGKRFQDPTTELAAKAAPLRPGY
FT E -> PRREDHLPNGVDQFQWHTKPSLPKPKANKEKTAR
FT RDGTCL (in isoform 4).
FT /FTID=VSP_000727.
FT Missing (in isoform 4).
FT /FTID=VSP_000728.
FT MISSING (IN REF. 2).
FT L -> S (IN REF. 2).
FT E -> K (IN REF. 2).
FT E -> K (IN REF. 2).
FT G -> S (IN REF. 1).
FT S
SQ SEQUENCE 2564 AA; 288982 MW; 52CDE7D11D601ECC CRC64;
Query Match 14.2%; Score 68; DB 1; Length 2564;
Best Local Similarity 26.2%; Pred. No. 1.4e+02;
Matches 21; Conservative 16; Mismatches 29; Indels 14; Gaps 4;
QY 5 QWNFAGIEAASAIQGNVTSIHSLDDEGKSLTKLAAAGSGSGSEAYQGVQ---KWDAT 61
Db 1021 QWRLSGIEALQALEPQA---ALLBEA---ALLAEFPQAARLHQGAELGAEWGAL 1073
QY 62 ATELNALQNLARTISEAQ 81
Db 1074 A-----SAAQACGEAAVAAAR 1089
RESULT 23
FLIC_SHIFL STANDARD; PRT; 550 AA.
ID FLIC_SHIFL
AC Q08860;

[illegible]

```

FT REPEAT      1032   1037       12.
REPEAT      1043   1048       13.
FT REPEAT      1052   1057       14.
REPEAT      1068   1063       15.
SQ SEQUENCE    1063 AA; 115649 MW; E4324DDBG1E4721A CRC64;

Query Match          14.0%; Score 67; DB 1; Length 1063;
Best Local Similarity 21.9%; Pred. No. 65;
Matches 28; Conservative 9; Mismatches 43; Indels 48; Gaps 3;

QY           9 AGTAAASAIOGNVTSIHSLDEGGKOSLYKL----- 39
Db         891 ACAAATSSGLSGMTFGWSSFDDGTTPAYNAHGGSGGVGSWGGAWSVGCGNGGCASAW 950
QY        40 -----AAAAGGSGSEAYQ-----GYQQKWDAIAT-----ELNNALQNLRITISEAG 80
Db       951 GGACGCCASANGGOCTCATSTWGCASAWNKSWWGGASTWASGESNGANMSTWGGTGDRSA 1010
QY            81 QAMASTEg 88
Db        1011 YGGCAStwG 1018
```

Search completed: February 5, 2004, 17:38:48
Job time : 7.52381 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:34:22 ; Search time 13.7976 Seconds
(without alignments)
1776.756 Million cell updates/sec

Title: US-09-805-427A-1
Perfect score: 479
Sequence: 1 MTEQQWNPAGTEAAASAIQGG.....ISEAGQWASTEGNVTGMFA 95

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830521

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 90 summaries

Database :

- SPTREMBL_23:*
- 1: sp archaea:*
- 2: sp bacteria:*
- 3: sp fungi:*
- 4: sp human:*
- 5: sp invertebrate:*
- 6: sp mammal:*
- 7: sp mhc:*
- 8: sp organelle:*
- 9: sp phage:*
- 10: sp plant:*
- 11: sp rodent:*
- 12: sp virus:*
- 13: sp vertebrate:*
- 14: sp unclassified:*
- 15: sp rvirus:*
- 16: sp bacteriaph:*
- 17: sp archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	151	31.5	96	16 Q8NST7	Q8nst7 corynebacte
2	150	31.3	96	16 Q8FS23	Q8fs23 corynebacte
3	103	21.5	96	16 Q8G878	Q8g878 bifidobacte
4	91	19.0	1460	16 Q8K6B1	Q8k6b1 streptococ
5	86.5	18.1	100	16 Q8G261	Q8g261 mycobacteri
6	84	17.5	142	10 Q8H0Q4	Q8h0q4 hordium vul
7	81.5	17.0	108	16 Q8NST8	Q8nst8 corynebacte
8	79	16.5	104	16 Q9CCV7	Q9ccv7 mycobacteri
9	77.5	16.2	135	16 Q8G262	Q8g262 mycobacteri
10	77	16.1	876	16 Q8G727	Q8g727 bifidobacte
11	76.5	16.0	903	2 Q8VTF1	Q8vtf1 bacillus st
12	76.5	16.0	9904	16 Q8NWQ6	Q8nwq6 staphylococ
13	75.5	15.8	106	16 Q8FS24	Q8fs24 corynebacte
14	75.5	15.8	305	10 Q9LLS8	Q9lls8 avena sativ
15	75	15.7	490	16 Q8HXZ9	Q8hxx9 pseudomonas
16	75	15.7	1037	16 Q8XEh2	Q8xeh2 escherichia

17	75	15.7	1037	16 Q8FF84	Q8ff84 escherichia
18	74.5	15.6	186	5 Q17179	Q17179 bombyx mori
19	73.5	15.3	186	5 Q17180	Q17180 bombyx mori
20	73.5	15.3	331	10 Q8S9Q2	Q8s9q2 oryza sativ
21	73.5	15.3	371	16 Q97M89	Q97m89 clostridium
22	73.5	15.3	389	2 Q9ZIX5	Q9zix5 streptococ
23	73.5	15.3	772	16 Q8E2M9	Q8e2m9 streptococ
24	73.5	15.3	772	16 Q8DWS9	Q8dws9 streptococ
25	73.5	15.3	2785	16 Q8EXK6	Q8exk6 streptococ
26	73.5	15.3	3141	16 Q8F850	Q8f850 leptospira
27	73	15.2	760	5 Q96H0H	Q96h0h drosophila
28	73	15.2	1204	16 Q915S5	Q915s5 pseudomonas
29	73	15.2	2470	5 Q9VK45	Q9vk45 drosophila
30	72.5	15.1	238	10 Q49013	Q49013 glycine max
31	72.5	15.1	344	12 Q91BA2	Q91ba2 spodoptera
32	72.5	15.1	501	13 Q91BD0	Q91bd0 brachydanio
33	72.5	15.1	756	2 Q44102	Q44102 anaplasma m
34	72	15.0	96	16 Q9KE84	Q9ke84 bacillus ha
35	72	15.0	637	16 Q9AAK8	Q9aa84 caulobacter
36	72	15.0	1225	16 Q9RZ46	Q9rz46 deinococcus
37	72	15.0	4631	4 Q8WY2	Q8wxy2 homo sapien
38	72	15.0	5938	4 Q96PK2	Q96pk2 homo sapien
39	72	15.0	6713	16 Q99U54	Q99u54 staphylococ
40	72	15.0	6713	16 Q931R6	Q931r6 staphylococ
41	71.5	14.9	526	16 Q9RUB1	Q9rub1 deinococcus
42	71.5	14.9	927	16 Q928S2	Q928s2 listeria in
43	71	14.8	97	16 Q97M29	Q97m29 clostridium
44	71	14.8	185	13 Q8UJ68	Q8uj68 oncorhynchu
45	71	14.8	334	10 Q9XEN7	Q9xen7 triticum ae
46	71	14.8	415	2 P95813	P95813 streptococ
47	71	14.8	896	16 Q8PAJ6	Q8paj6 xanthomonas
48	70.5	14.7	425	2 Q54860	Q54860 streptococ
49	70.5	14.7	485	9 Q8SCI0	Q8scio propionibac
50	70.5	14.7	656	16 Q9WYR0	Q9wyr0 thermotoga
51	70.5	14.7	656	16 Q9WXN0	Q9wxn0 thermotoga
52	70.5	14.7	661	16 Q9X020	Q9x020 thermotoga
53	70.5	14.7	957	10 Q94GY5	Q94gy5 oryza sativ
54	70.5	14.7	1510	9 Q80179	Q80179 streptococ
55	70.5	14.7	2475	2 Q82833	Q82833 bacillus sp
56	70	14.6	265	16 Q8NNM8	Q8nnm8 corynebacte
57	70	14.6	653	16 Q8FQL5	Q8fq15 corynebacte
58	70	14.6	710	13 Q91402	Q91402 xenopus. he
59	70	14.6	1339	5 Q9VW06	Q9vw06 drosophila
60	69.5	14.5	362	16 Q9F321	Q9f321 streptomyce
61	69.5	14.5	569	16 Q8XR59	Q8xr59 ralstonia s
62	69.5	14.5	664	16 Q8RBV6	Q8rbv6 thermoanaer
63	69.5	14.5	742	2 Q9X4H8	Q9x4h8 anaplasma m
64	69.5	14.5	825	17 Q8TJH9	Q8tjh9 methanosarc
65	69.5	14.5	1024	10 Q9SKV4	Q9skv4 arabidopsis
66	69.5	14.5	1509	16 Q9SST6	Q9sst6 staphylococ
67	69.5	14.5	1509	16 Q931M0	Q931m0 staphylococ
68	69.5	14.5	3165	16 Q8RDQ9	Q8rdq9 fuscobacteri
69	69	14.4	205	9 Q8LTT0	Q8ltt0 bacterioph
70	69	14.4	221	16 Q8NM43	Q8nm43 corynebacte
71	69	14.4	222	2 Q93EL3	Q93el3 borrelia bu
72	69	14.4	337	17 Q972P4	Q972p4 sulfolobus
73	69	14.4	472	2 Q8GQ61	Q8gqe1 pseudomonas
74	69	14.4	484	4 Q96RQ5	Q96rq5 homo sapien
75	69	14.4	484	11 Q9R0B2	Q9r0b2 mus musculu
76	69	14.4	484	11 Q9Z187	Q9z187 mus musculu
77	69	14.4	488	4 Q96NW9	Q96nw9 homo sapien
78	69	14.4	490	4 Q9UL67	Q9ul67 homo sapien
79	69	14.4	508	4 Q95319	Q95319 homo sapien
80	69	14.4	508	11 Q9Z0H4	Q9z0h4 mus musculu
81	69	14.4	509	4 Q96RQ6	Q96rq6 homo sapien
82	69	14.4	509	4 Q92950	Q92950 homo sapien
83	69	14.4	521	4 Q8N499	Q8n499 homo sapien
84	69	14.4	532	11 Q88756	Q88756 rattus norv
85	69	14.4	572	16 Q92TP8	Q92tp8 rhizobium m
86	69	14.4	1037	16 Q8ZN77	Q8zn77 salmonella
87	69	14.4	1037	16 Q8Z4S4	Q8z4s4 salmonella
88	68.5	14.3	103	16 Q86644	Q86644 streptomyce
89	68.5	14.3	138	10 Q8S386	Q8s386 secale cere

90	68.5	14.3	414	10	Q8RZT1	Q8rzt1 oryza sativ
ALIGNMENTS						
RESULT 1						
Q8NST7	PRELIMINARY; PRT; 96 AA.					
AC	Q8NST7;					
DT	01-OCT-2002 (TrEMBLrel. 22, Created)					
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)					
DT	01-OCT-2002 (TrEMBLrel. 22, Last annotation update)					
DE	Hypothetical protein Cgl0580.					
GN	Cgl0580.					
OS	Corynebacterium glutamicum (Brevibacterium flavum).					
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;					
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.					
OX	NCBI_TaxID=1718;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;					
RA	Nakagawa S.;					
RT	"Complete genomic sequence of Corynebacterium glutamicum ATCC 13032.";					
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.					
DR	EMBL; AF005275; BAB97973.1; -.					
KW	Hypothetical protein; Complete proteome.					
SQ	SEQUENCE 96 AA; 10324 MW; 724318FEAE03AD76 CRC64;					
Query Match 31.5%; Score 151; DB 16; Length 96;						
Best Local Similarity 37.5%; Pred. No. 4.7e-07;						
Matches 30; Conservative 14; Mismatches 36; Indels 0; Gaps 0;						
QY	5	QWVFAGTAAASAIQGNVTSHSLDEGKQSLTKLAAWGGSGSEAYQGVQKWDATATE	64			
DB	5	RYEFGAIQGAATDINSTGRINSLDGLKSQLQPMVATWEGDSAYAYNEAQKWDRAAAE	64			
QY	65	LNALQNLARTISEAQGAMA	84			
DB	65	LNLTILATISNTVAQGAERMS	84			
RESULT 2						
Q8FS23	PRELIMINARY; PRT; 96 AA.					
ID	Q8FS23;					
AC	Q8FS23;					
DT	01-MAR-2003 (TrEMBLrel. 23, Created)					
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)					
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)					
DE	Conserved hypothetical protein.					
GN	CE0583.					
OS	Corynebacterium efficiens.					
OC	Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;					
OC	Corynebacterineae; Corynebacteriaceae; Corynebacterium.					
OX	NCBI_TaxID=152794;					
RN	[1]					
RP	SEQUENCE FROM N.A.					
RC	STRAIN=YS-314 / AJ 12310 / DSM 44549 / JCM 11189;					
RA	Kawarayashi Y., Yamazaki J., Hino Y., Kikuchi H., Nakamura Y.,					
RA	Ikeo K., Suzuki M., Mashima J., Itoh T., Yamagishi A., Nishio Y.,					
RA	Usuda Y., Sugimoto S.;					
RT	"The entire genomic sequence of Corynebacterium efficiens YS-314.";					
RL	Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.					
DR	EMBL; AF005215; BAC17393.1; -.					
KW	Hypothetical protein; Complete proteome.					
SQ	SEQUENCE 96 AA; 10310 MW; EF8D4BDC9E4F0A5 CRC64;					
Query Match 31.3%; Score 150; DB 16; Length 96;						
Best Local Similarity 37.5%; Pred. No. 5.9e-07;						
Matches 30; Conservative 14; Mismatches 36; Indels 0; Gaps 0;						
QY	5	QWVFAGTAAASAIQGNVTSHSLDEGKQSLTKLAAWGGSGSEAYQGVQKWDATATE	64			

Db	5	RYEFGAIQGAATDINSTGRINSLDGLKSQLQPMVATWEGDSAYAYNEAQKWDNAAAAE	64
QY	65	LNALQNLARTISEAQGAMA	84
Db	65	LNLTILATISNTVAQGAERMS	84
RESULT 3			
Q8G878	PRELIMINARY; PRT; 96 AA.		
ID	Q8G878;		
AC	Q8G878;		
DT	01-MAR-2003 (TrEMBLrel. 23, Created)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Hypothetical protein.		
GN	BL0033.		
OS	Bifidobacterium longum.		
OC	Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;		
OC	Bifidobacteriaceae; Bifidobacterium.		
OX	NCBI_TaxID=216816;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=NCC 2705;		
RA	MEDLINE=22294977; PubMed=12381787;		
RA	Schell M.A., Karimantzou M., Shet B., Vilanova D., Berger B.,		
RA	Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,		
RA	Pridmore R.D., Arigoni F.;		
RT	"The genome sequence of Bifidobacterium longum reflects its adaptation to the human gastrointestinal tract.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427(2002).		
DR	EMBL; AE014614; AA023870.1; -.		
KW	Hypothetical protein; Complete proteome.		
SQ	SEQUENCE 96 AA; 10431 MW; 7D02BE38C6B1F61C CRC64;		
Query Match 21.5%; Score 103; DB 16; Length 96;			
Best Local Similarity 22.1%; Pred. No. 0.018;			
Matches 21; Conservative 29; Mismatches 45; Indels 0; Gaps 0;			
QY	1	MTEQWNFAGIEAAASAIQGNVTSHSLDEGKQSLTKLAAWGGSGSEAYQGVQKWDKA	60
DB	1	MPQYQVDSERIQSSAAVATISIQIRQAVGMYTNALQDANRGSATQFTVAEQWRA	60
QY	61	TATELNALQNLARTISEAQGAMASTEIGNVTGMPA	95
DB	61	ACQQMEASLESIQRSLSLTQASTVYADAEIQASRLFA	95
RESULT 4			
Q8K6B1	PRELIMINARY; PRT; 1460 AA.		
ID	Q8K6B1;		
AC	Q8K6B1;		
DT	01-OCT-2002 (TrEMBLrel. 22, Created)		
DT	01-OCT-2002 (TrEMBLrel. 22, Last sequence update)		
DT	01-MAR-2003 (TrEMBLrel. 23, Last annotation update)		
DE	Putative tail protein-phage associated.		
GN	SPYM3 1421.		
OS	Streptococcus pyogenes (serotype M3).		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;		
OC	Streptococcus.		
OX	NCBI_TaxID=198466;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MGAS315 / Serotype M3;		
RX	MEDLINE=22133808; PubMed=12122206;		
RA	Beres S.B., Sylva G.D., Barbian K.D., Lei B., Hoff J.S.,		
RA	Mammarella N.D., Liu M.-Y., Smoot J.C., Porcella S.F., Parkins L.D.,		
RA	Campbell D.S., Smith T.M., McCormick J.K., Leung D.Y.M.,		
RA	Schlievert P.M., Musser J.M.;		
RT	"Genome sequence of a serotype M3 strain of group A Streptococcus: phage-encoded toxins, the high-virulence phenotype, and clone emergence.";		
RL	Proc. Natl. Acad. Sci. U.S.A. 99:10078-10083(2002).		
DR	EMBL; AE014163; AAM80028.1; -.		

DR InterPro; IPR001680; WD40.
 DR PROSITE; PS00678; WD REPEATS 1; 1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 1460 AA; 149590 MW; 109AB8DDFBD4A4B8 CRC64;

Query Match 19.0%; Score 91; DB 16; Length 1460;
 Best Local Similarity 33.0%; Pred. No. 7.5;
 Matches 31; Conservative 13; Mismatches 42; Indels 8; Gaps 4;

QY 4 CQWN--FAGIEAASAIQGNVTSHLSLDEGKQSLTKLAAWGGSGSEAYQGVQKWDAT 61
 DB 222 QWNFFSVSAKAFANKVST--SLFGLRTSLANLSKIGSSLS--NGFSRMSSSA 276
 QY 62 ATELNNALQNARTISEAGQMASTEGNVTGMFA 95
 DB 277 ATSLNGISQKFANT--SSAGERLKSIVMSIQVAFS 309

RESULT 5
 O06261
 ID O06261 PRELIMINARY; PRT; 100 AA.
 AC O06261;
 DT 01-JUL-1997 (TReMBLrel. 04, Created)
 DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
 DE Hypothetical protein RV3444C.
 GN RV3444C OR M7CY77.16C OR M73549.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala E.,
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RA "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544 (1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OshKosh;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 Peterson J., DeBoy R., Dodson R., Winn M.L., Haft D., Hickey E.,
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 Delcher A., Uterback T., Weidman J., Khouri H., Gill J., Mikula A.,
 Bishai W.;
 RA "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains."
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z95389; CAB08674.1;
 DR EMBL; AE007159; AKA47890.1;
 DR TIGR; MT3549;
 DR TubercuList; RV3444C;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 100 AA; 11120 MW; 5045B1B46578DC6C CRC64;

Query Match 18.1%; Score 86.5; DB 16; Length 100;
 Best Local Similarity 24.2%; Pred. No. 0.7;
 Matches 22; Conservative 13; Mismatches 55; Indels 1; Gaps 1;

QY 6 WNPAGIE-AAASAIQGNVTSHLSLDEGKQSLTKLAAWGGSGSEAYQGVQKWDATATE 64
 DB 9 YNFDALEYVROSHHTAARFNAALQELRSQAPLQQLWTREAAAYHAEQLKWHQAASA 68
 QY 65 LNNALQNARTISEAGQMASTEGNVTGMFA 95

Db 69 LNEILLDLGNVVRHGADDDVAHADRAAGAWA 99

RESULT 6
 Q8HQ04
 ID Q8HQ04 PRELIMINARY; PRT; 142 AA.
 AC Q8HQ04;
 DT 01-MAR-2003 (TReMBLrel. 23, Created)
 DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
 DE Hypothetical protein precursor.
 GN COR14B.
 OS Hordeum vulgare (Barley).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
 OC Triticeae; Hordeum.
 OX NCBI_TaxID=4513;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=cv. Aurea;
 RA Dal Bosco C., Busconi M., Govoni C., Baldi P., Stanca M.A.,
 Crosatti C., Bassi R., Cattivelli L.;
 RT "Cor gene expression in barley mutants affected in chloroplast
 development and photosynthetic electron transport."
 RL Plant Physiol. 0:0-0(0).
 DR EMBL; AJ512944; CAD55692.1;
 KW Transit peptide; Chloroplast.
 FT TRANSIT 1 52 CHLOROPLAST.
 FT CHAIN 53 142 COLD REGULATED PROTEIN.
 SQ SEQUENCE 142 AA; 13919 MW; EC4860B010D59694 CRC64;

Query Match 17.5%; Score 84; DB 10; Length 142;
 Best Local Similarity 22.9%; Pred. No. 1.9;
 Matches 16; Conservative 19; Mismatches 35; Indels 0; Gaps 0;

QY 12 EAAASAIQGNVTSHLSLDEGKQSLTKLAAWGGSGSEAYQGVQKWDATATELNNALQN 71
 DB 73 DATDNVAGAKSAGESVVDAAKQASGVSETAQLDGGQAKQAEAWDATKDAASQVADN 132
 QY 72 LARTISEAGQ 81
 DB 133 VAAAVEDVSK 142

RESULT 7
 Q8NST8
 ID Q8NST8 PRELIMINARY; PRT; 108 AA.
 AC Q8NST8;
 DT 01-OCT-2002 (TReMBLrel. 22, Created)
 DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
 DE Hypothetical protein Cgl0579.
 GN CGL0579.
 OS Corynebacterium glutamicum (Brevibacterium flavum).
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=1718;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 13032 / DSM 20300 / NCIB 10025;
 RA Nakagawa S.;
 RT "Complete genomic sequence of Corynebacterium glutamicum ATCC 13032."
 RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP005275; BAB97972.1;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 108 AA; 11644 MW; E0B3AFA03AE9D452 CRC64;

Query Match 17.0%; Score 81.5; DB 16; Length 108;
 Best Local Similarity 18.2%; Pred. No. 2.3;
 Matches 14; Conservative 24; Mismatches 32; Indels 7; Gaps 1;

QY 18 IQGNVTSHLSLDEGKQSLTKLAAWGGSGSEAYQGVQKWDATATELNNALQNARTIS 77

```

Ra Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
Ra Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,
Ra Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy J.,
Ra Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
Ra Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
Ra Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence.";
RL Nature 393:537-544 (1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
Ra Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
Ra Peterson J., DeBoy R., Dodson R., Gwinn M., Haft D., Hickey E.,
Ra Klotz J.F., Nelson W.C., Umayam L.A., Ermolaeva M., Salzberg S.L.,
Ra Deichner A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
Ra Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z95389; CAB08675.1; ALT_INIT.
DR EMBL; AE007159; AAK47891.1; -.
DR TIGR; MT3550; -.
DR Tuberculist; RV3445C; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 135 AA; 14592 MW; A6BC5AED45621C07 CRC64;

Query Match 16.2%; Score 77.5; DB 16; Length 135;
Best Local Similarity 35.3%; Pred. No. 7.4; 18; Indels 7; Gaps 1;
Matches 18; Conservative 8; Mismatches 18;

QY 41 AAWGSGSEAYGVQVQKWDATATLNNALQNARTI-----SEAGQAMA 84
Db 75 SVWGGLAARFQDVDRWNAESTRLVHLHAADTIRHNEAALREAGQIHA 125

RESULT 10
Q8G727 PRELIMINARY; PRT; 876 AA.
AC Q8G727;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Hypothetical membrane protein with similarity to phage infection
DE protein.
GN BL0446.
OS Bifidobacterium longum.
OC Bacteria; Actinobacteria; Actinobacteridae; Bifidobacteriales;
OC Bifidobacteriaceae; Bifidobacterium.
OX NCBI_TaxID=216816;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NCC 2705;
RX MEDLINE=22294977; PubMed=12381787;
Ra Schell M.A., Karmirantzou M., Snel B., Villanova D., Berger B.,
Ra Pessi G., Zwaalen M.-C., Desiere F., Bork P., Delley M.,
Ra Bridmore R.D., Arigoni F.;
RT "The genome sequence of Bifidobacterium longum reflects its adaptation
RT to the human gastrointestinal tract.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:14422-14427 (2002).
DR EMBL; AE014664; AAN24282.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 876 AA; 93283 MW; A5966B73BF878545 CRC64;

Query Match 16.1%; Score 77; DB 16; Length 876;
Best Local Similarity 26.7%; Pred. No. 86;
Matches 23; Conservative 18; Mismatches 33; Indels 12; Gaps 2;

QY 11 IEAASAIQGNVTSIHSLLDEGKQSLTKLAAAGGS-----GSEAYGVQVQKWDAT 64
Db 412 LSSSTASTLSTGITSQGLIQSKHALDQDKAASTVTALRDTDKALAGVQTKLDTLVT 471
QY 65 LNNALQNARTISEAGQAMASTEAGNV 90

```


"A kernel-specific 1,3-beta glucanase in oat."; Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

```
DR HSP; P15737; IGHS.
DR InterPro; IPR000490; Glyco_hydro_17.
DR Pfam; PF00332; Glyco_hydro_17; 1.
FT NON_TER
SQ SEQUENCE 305 AA; 32358 MW; FEE831E77D0872EE CRC64;

Query Match 15.8%; Score 75.5; DB 10; Length 305;
Best Local Similarity 25.6%; Pred. No. 32;
Matches 31; Conservative 22; Mismatches 39; Indels 29; Gaps 5;

QY 1 MTEQOWNFAGLEAAASALQGNVTSIHSLDEGKQSLTKL-----AAAGSGSEAYQGV 54
DB 27 ITDMRIYFADQ-ALSALRN--SGILVMDIGNDQKGIAGSANAARWFDNVQRQGL 83
QY 55 QOKWDATATELNN-----ALQNLARTISSEAG-----QAMASTEGNVTGMF 94
DB 84 KIKIVAGNEIQGGDTGRIVPAIRNLNLAALSAAGLIGIKVSTAIRFDVANSFPESAGVF 143
QY 95 A 95
DB 144 A 144

RESULT 15
Q9H2X9 PRELIMINARY; PRT; 490 AA.
AC Q9H2X9;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 23, Last annotation update)
DE Probable chemotaxis transducer.
GN PA2867.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PAO1;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
RT opportunistic pathogen.";
RL Nature 406:959-964 (2000).
DR EMBL; AE004713; AAG06255.1; -.
DR HSP; P02942; 1QJ7.
DR InterPro; IPR004089; Chmtaxis_transd.
DR Pfam; PF00672; HAMP; 1.
DR Pfam; PF00015; MCPsignal; 1.
DR SMART; SM00304; HAMP; 1.
DR SMART; SM00283; MA; 1.
DR PROSITE; PS01111; CHEMOTAXIS_TRANSLOC_2; 1.
KW Complete proteome.
SQ SEQUENCE 490 AA; 51457 MW; 1ED660B984A963CA CRC64;

Query Match 15.7%; Score 75; DB 16; Length 490;
Best Local Similarity 29.1%; Pred. No. 64;
Matches 23; Conservative 12; Mismatches 36; Indels 8; Gaps 2;

QY 8 FAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAGSGSEAYQGVQOKWDATATELNN 67
DB 201 FNAWQGYQRIYVTVAAATKLDGAQALARM-----GQVRQGMIGQSEETDQTAIN- 255
QY 68 ALQNLARTISSEAGQAMST 86
DB 256 ---EMSTTVPHIAQHAAT 271
```

```
RESULT 16
Q8XE2 PRELIMINARY; PRT; 1037 AA.
AC Q8XE2;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Sensitivity to acriflavine, integral membrane protein, possible efflux
DE pump (Aminoglycoside efflux pump) (Putative transport system permease
DE protein).
DE ACRD OR Z3727 OR ECS3332.
GN Escherichia coli O157:H7.
OS Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Escherichia.
OX NCBI_TaxID=8334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Roca D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamocous K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Iida T., Takami H., Honda T., Sasaki K., Murata T., Tanaka M., Tohe T.,
RA Kubara S., Shiba T., Hattori M., Shinagawa H.;
RA "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22 (2001)
DR EMBL; AE005476; AAG57579.1; -.
DR EMBL; AP002561; BAB36755.1; -.
DR InterPro; IPR001036; Acriflavin_res.
DR InterPro; IPR004764; HAE1.
DR Pfam; PF00873; ACR tran; 1.
DR PRINTS; PR00702; ACRIFLAVINRP.
DR TIGRFAMs; TIGR00915; 2A0602; 1.
KW Complete proteome.
FT CONFLICT 127 127 V -> L (IN REF. 2).
SQ SEQUENCE 1037 AA; 113019 MW; B44742633B1DA058 CRC64;

Query Match 15.7%; Score 75; DB 16; Length 1037;
Best Local Similarity 26.2%; Pred. No. 1.6e+02;
Matches 27; Conservative 18; Mismatches 39; Indels 20; Gaps 4;

QY 7 NFAGIEA-----AASAIQGNVTSIHSL-----DEGKQSLTKLAAGSGSEAYQGV 53
DB 48 NYPGASQULENTVTVQVIEQNTGLNLMYSSQSGTGQASVT-LSFKAGTDPDEAVQV 106
QY 54 VQOKWDATATELNNALQNLARTISAGQ-----AMASTEGNV 90
DB 107 VQNLQSAARKLPQAVQNGQVTVRKGTGNTNLTIAFVSTDGS 149

RESULT 17
Q8FF84 PRELIMINARY; PRT; 1037 AA.
AC Q8FF84;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Probable aminoglycoside efflux pump.
GN ACRD OR C2997.
OS Escherichia coli O6.
```

OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Escherichia.
 OX NCBI_TaxID=217992;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=O6.H1 / CFT073 / ATCC 700928;
 RX MEDLINE=22388234; PubMed=12471157;
 RA Welch R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
 RA Rasko D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
 RA Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
 RA Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
 RT "Extensive mosaic structure revealed by the complete genome sequence
 of uropathogenic *Escherichia coli*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024(2002).
 DR EMBL; AE016764; AAN81447.1; --
 KW Complete proteome.
 SQ SEQUENCE 1037 AA; 113093 MW; 94CE259053DF1461 CRC64;

 Query Match 15.7%; Score 75; DB 16; Length 1037;
 Best Local Similarity 26.2%; Pred. No. 1.6e+02;
 Matches 27; Conservative 18; Mismatches 38; Indels 20; Gaps 4;

 QY 7 NFAGLEA-----AASATQGNVTSIHSLL-----DEGKSLTKLAAAWGSGSEAYQG 53
 Db 48 NYPGASQATLENTVTQVIEQNTGLDNLMYMSQSSGTGQASVT-LSFKAGTDPDEAVQQ 106
 QY 54 VQOKWDATATELNALQNLARTISEAQ-----AMASTEQNV 90
 Db 107 VQNLQSNMRKLPQAVQNGVTVRKTGDTNLTITAFVSTDGSM 149

 PRELIMINARY; PRT; 186 AA.

 Q17179
 AC Q17179;
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)
 DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
 DE Apolipoprotein-III precursor.
 OS Bombyx mori (Silk moth).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Bombycoidea;
 OC Bombycidae; Bombyx.
 OX NCBI_TaxID=7091;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=N4; TISSUE=Fat body;
 RX MEDLINE=20082772; PubMed=10613959;
 RA Yamauchi Y., Hooffer C., Yamamoto A., Takeda H., Ishihara R.,
 RA Maekawa H., Sato R., Su-Ii S., Sumida M., Wells M.A., Tsuchida K.;
 RT "cDNA and deduced amino acid sequences of apolipoprotein-III from
 Bombyx mori and Bombyx mandarina.";
 RL Arch. Insect Biochem. Physiol. 43:16-21 (2000).
 DR EMBL; U59243; AAB02852.1; --
 KW Signal.
 FT SIGNAL
 FT CHAIN
 SQ SEQUENCE 186 AA; 20771 MW; B5F66548C439D1BC CRC64;

 Query Match 15.6%; Score 74.5; DB 5; Length 186;
 Best Local Similarity 27.7%; Pred. No. 21;
 Matches 33; Conservative 11; Mismatches 38; Indels 37; Gaps 5;

 QY 8 FAGIEAASA-----IQGNVTSIHSLLDEGKSLTK-----LAAAWGSGSE 49
 Db 9 FACISLAQGANVRDAPDFFKDIEHTKTFHKTLEQQFNLSLTKSDAQDFSKAW-KDGS 67
 QY 50 AYGVQVQKWDATATELNAL-----QNLARTISEAGQAMASTEQNVGM 93
 Db 68 S---VLQQLNFAKSLQGLGDANAKAKALEQSRNIERTAEILRKAHPDVEKNATAL 123

Query Match 15.3%; Score 73.5; DB 10; Length 331;
Best Local Similarity 32.9%; Pred. No. 55;
Matches 26; Conservative 13; Mismatches 27; Indels 13; Gaps 4;
QY 13 AAASAIQGNVTSIHSLDEG--KSLTKLAAA-WGSGSEAYQGVQKWDATATLNN-- 67
DB 66 ATLRLAG--TJIAVDEPAIDQLTLTSAASDWQSNIKPYQGVNIYIAVGNVSGDA 123
QY 68 -----ALQNLRITISEAG 80
DB 124 TRESILPAMENLTAKLSAAG 142
RESULT 21
Q97M89 PRELIMINARY; PRT; 371 AA.
AC Q97M89;
DT 01-OCT-2001 (TREMELrel. 18, Created)
DT 01-OCT-2001 (TREMELrel. 18, Last sequence update)
DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
DE Cell wall-associated hydrolase.
GN CAC0309.
OS Clostridium acetobutylicum.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1488;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 824 / DSM 792 / VKM B-1787;
RX MEDLINE=21359325; PubMed=11466286;
RA Noelling J., Bretton G., Omeichenko M.V., Makarova K.S., Zeng Q.,
Gibson R., Lee H.M., Dubois J., Qiu D., Hitti J., Wolf Y.I.,
Tatusov R.L., Sabathe F., Doucette-Stamm L., Soucaille P., Daly M.J.,
Bennett G.N., Koonin E.V., Smith D.R.;
RA "Genome sequence and comparative analysis of the solvent-producing
bacterium Clostridium acetobutylicum";
RT J. Bacteriol. 183:4823-4838(2001).
RL EMBL; AEO07545; AAK78290.1; -
DR InterPro; IPR000064; NLP_C60.
DR Pfam; PF00877; NLP_C60; 1.
KW Hydrolase; Complete proteome.
SQ SEQUENCE 371 AA; 41319 MW; B817536899C7EB8A CRC64;
Query Match 15.3%; Score 73.5; DB 16; Length 371;
Best Local Similarity 28.6%; Pred. No. 63;
Matches 24; Conservative 17; Mismatches 32; Indels 11; Gaps 3;
QY 10 GIEAASAIQGNVTSIHSLDEGKSLTKLAAAWGSGSEAYQGVQKWDATAT---ELN 66
DB 13 GITVSCS---GNIVFASFLQDQYNQSQCYQNA-----LKSVDIENKIALDNQIGELN 64
QY 67 NALQNLRITISEAGQAMASTEGNV 90
DB 65 NSINDTKRINESQNNAITQGI 88
RESULT 22
Q9ZIX5 PRELIMINARY; PRT; 389 AA.
AC Q9ZIX5;
DT 01-MAY-1999 (TREMELrel. 10, Created)
DT 01-MAY-1999 (TREMELrel. 10, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Mip22.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1314;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AL168;
RA Thern A., Wustfelt M., Lindahl G.;
Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RL

-1- SUBCELLULAR LOCATION: ATTACHED TO THE CELL WALL PEPTIDOGLYCAN BY
AN AMIDE BOND (BY SIMILARITY).
EMBL; AF007557; AAD01409.1; -
DR InterPro; IPR005877; Gpos_YSIIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR006192; LPXTG.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF04650; YSIIRK signal; 1.
DR PRINTS; PR00015; GPOSANCHOR.
DR TIGRFAMs; TIGR01167; LPXTG anchor; 1.
DR TIGRFAMs; TIGR01168; YSIIRK signal; 1.
DR PROSITE; PS0847; GRAM_POS_ANCHORING; 1.
KW Cell wall; Peptidoglycan-anchor.
SQ SEQUENCE 389 AA; 42424 MW; 68F2E0CF15C5EC CRC64;
Query Match 15.3%; Score 73.5; DB 2; Length 389;
Best Local Similarity 24.4%; Pred. No. 67;
Matches 21; Conservative 19; Mismatches 41; Indels 5; Gaps 1;
QY 1 MTEQQNFAGIEAASAIQGNVTSIHSLDEGKSLTKLAAAWGSGSEAYQGVQKWDAT 60
DB 204 LTELQAKLDTATAEKAKLESRTVLTLENLGSARLTDLQAKLDASNAE-----KEKLOS 258
QY 61 TATELNALQNLRITISEAGQAMAST 86
DB 259 QAATLEKQLEATKKEKLADLQAKLAAT 284
RESULT 23
Q8E2M9 PRELIMINARY; PRT; 772 AA.
AC Q8E2M9;
DT 01-MAR-2003 (TREMELrel. 23, Created)
DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
DT 01-MAR-2003 (TREMELrel. 23, Last annotation update)
DE Hypothetical protein.
GN GBS2093.
OS Streptococcus agalactiae (serotype III).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216495;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NEM316 / Serotype III;
RX MEDLINE=22242508; PubMed=12354221;
RA Glaser P., Rusniok C., Buchrieser C., Chevallier F., Frangeul L.,
Meadek T., Zouine M., Couve E., Lalioui L., Poyart C., Trieu-Cuot P.,
Kunst F.;
RA "Genome sequence of Streptococcus agalactiae, a pathogen causing
invasive neonatal disease";
RL Mol. Microbiol. 45:1499-1513(2002).
DR EMBL; AL766856; CAD47752.1; -
DR Sagalish; gbs2093; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 772 AA; 82819 MW; 4F3B4A1IDE20A850 CRC64;
Query Match 15.3%; Score 73.5; DB 16; Length 772;
Best Local Similarity 26.2%; Pred. No. 1.6e+02;
Matches 22; Conservative 12; Mismatches 47; Indels 3; Gaps 2;
QY 10 GIEAASAIQGNVTSIHSLDEGKSLTKLAAAWGSGSEAYQGVQKWDAT--ATELNN 67
DB 238 GLETLSDGVATYTTGVHK-LSEGSQKLDKSAALVEGSEKLTDLGLOLSQLQPKPOER 296
QY 68 ALQNLRITISEAGQAMASTEGNV 91
DB 297 TLQNLSDGLKNLQIITNLQSTAT 320
RESULT 24
Q8DWS9 PRELIMINARY; PRT; 772 AA.
ID Q8DWS9
AC Q8DWS9;

```
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Membrane protein, putative.
GN SAG2134.
OS Streptococcus agalactiae (serotype V).
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=216466;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2603 V/R / Serotype V;
RX MEDLINE=22222988; PubMed=12200547;
RA Tettelin H., Masiugnani V., Cieslewicz M.J., Eisen J.A., Peterson S.,
RA Wessels M.R., Paulsen I.T., Nelson K.E., Margarit I., Read T.D.,
RA Madoff L.C., Wolf A.M., Beanan M.J., Brinkac L.M., Daugherty S.C.,
RA DeBoff R.T., Durkin A.S., Kolonay J.F., Maduri R., Lewis M.R.,
RA Radune D., Fedorova N.B., Scanlan D., Khouri H., Mulligan S.,
RA Carty H.A., Cline R.T., Van Aken S.E., Gill J., Scarselli M., Mora M.,
RA Iacobini E.T., Brettoni C., Galli G., Mariani M., Vegni F., Maione D.,
RA Rinaldo D., Rappuoli R., Telford J.L., Kasper D.L., Grandi G.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative genomic analysis of an
RT emerging human pathogen, serotype V Streptococcus agalactiae.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:12391-12396 (2002).
DR EMBL; AE014288; AAN00992.1; -
DR TIGR; SAG2134; -
KW Complete proteome.
SQ SEQUENCE 772 AA; 82807 MW; D4B027A703DCA790 CRC64;

Query Match 15.3%; Score 73.5; DB 16; Length 772;
Best Local Similarity 26.2%; Pred. No. 1.6e+02;
Matches 22; Conservative 12; Mismatches 4; Indels 3; Gaps 2;

QY 10 GIEAASAIQGNVTSIHSLLEGKQSLTKLAAAGSGSEAYQGVQKWDAT--ATELNN 67
Db 238 GLETLSDGTAYTTGVHK-LSEGSQKLDKQALVEGSEKLTGDI-QQLSQATQLKPEQR 296
QY 68 ALQNLARTISSEAGAMASTEAGNTV 91
Db 297 TLQNLSDGLKMLNQLITNLQSTAT 320

RESULT 25
Q8EZX6 PRELIMINARY; PRT; 2785 AA.
AC Q8EZX6;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Probable phenazine biosynthesis family protein.
GN LA3725.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011528; AAN50923.1; -
KW Complete proteome.
SQ SEQUENCE 2785 AA; 305026 MW; 743BB099DD3C4930 CRC64;

Query Match 15.3%; Score 73.5; DB 16; Length 2785;
Best Local Similarity 25.2%; Pred. No. 7.9e+02;
Matches 31; Conservative 21; Mismatches 32; Indels 39; Gaps 8;

QY 3 EQQWNFAGIEAASAIQGNVTSIH-----SLLDEGKQSLTKLAAAGG-----SGS 48
Db 164 QSQWD-----SFVSGNGLNLMNGSLSQTLVNSGQTLGLEGGQWVDFNNLQSG 215
QY 49 EAYQ----GVQKWDATATLNNALQNLARTI-----SEAG---QAMASTEAG-----NVT 91

RESULT 26
Q8F850 PRELIMINARY; PRT; 3141 AA.
AC Q8F850;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN LA0709.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011258; AAN47908.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 3141 AA; 343535 MW; B931EDC3DA2915F CRC64;

Query Match 15.3%; Score 73.5; DB 16; Length 3141;
Best Local Similarity 25.2%; Pred. No. 9.1e+02;
Matches 31; Conservative 21; Mismatches 32; Indels 39; Gaps 8;

QY 3 EQQWNFAGIEAASAIQGNVTSIH-----SLLDEGKQSLTKLAAAGG-----SGS 48
Db 166 QSQWD-----SFVSGNGLNLMNGSLSQTLVNSGQTLGLEGGQWVDFNNLQSG 217
QY 49 EAYQ----GVQKWDATATLNNALQNLARTI-----SEAG---QAMASTEAG-----NVT 91
Db 218 QTYQQAAGLTKYQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNL 277
QY 92 GMF 94
Db 278 GLF 280

RESULT 27
Q960H0 PRELIMINARY; PRT; 760 AA.
AC Q960H0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SD02289p.
GN TOR OR CG5092.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY052068; AAK93492.1; -
DR FlyBase; FBgn0021796; Tor.
DR InterPro; IPR003151; FAT.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000403; PI3_P14_kinase.
```

```
Db 216 QTYQQAAGLTKYQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNL 275
QY 92 GMF 94
Db 276 GLF 278

RESULT 26
Q8F850 PRELIMINARY; PRT; 3141 AA.
AC Q8F850;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DE Hypothetical protein.
GN LA0709.
OS Leptospira interrogans.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiroaceae; Leptospira.
OX NCBI_TaxID=173;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=56601 / Serogroup Icterohaemorrhagiae / Serovar lai;
RA Ren S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE011258; AAN47908.1; -
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 3141 AA; 343535 MW; B931EDC3DA2915F CRC64;

Query Match 15.3%; Score 73.5; DB 16; Length 3141;
Best Local Similarity 25.2%; Pred. No. 9.1e+02;
Matches 31; Conservative 21; Mismatches 32; Indels 39; Gaps 8;

QY 3 EQQWNFAGIEAASAIQGNVTSIH-----SLLDEGKQSLTKLAAAGG-----SGS 48
Db 166 QSQWD-----SFVSGNGLNLMNGSLSQTLVNSGQTLGLEGGQWVDFNNLQSG 217
QY 49 EAYQ----GVQKWDATATLNNALQNLARTI-----SEAG---QAMASTEAG-----NVT 91
Db 218 QTYQQAAGLTKYQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNLQNL 277
QY 92 GMF 94
Db 278 GLF 280

RESULT 27
Q960H0 PRELIMINARY; PRT; 760 AA.
AC Q960H0;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE SD02289p.
GN TOR OR CG5092.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise B., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celniker S.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY052068; AAK93492.1; -
DR FlyBase; FBgn0021796; Tor.
DR InterPro; IPR003151; FAT.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000403; PI3_P14_kinase.
```

```

DR Pfam; PF02259; FAT; 1.
DR Pfam; PF02260; FATC; 1.
DR Pfam; PF00454; P13_P14_kinase; 1.
DR SMART; SM00146; P13KC; 1.
DR PROSITE; PS00915; P13_4_KINASE_1; 1.
DR PROSITE; PS00916; P13_4_KINASE_2; 1.
DR PROSITE; PS00290; P13_4_KINASE_3; 1.
DR PROSITE; PS00290; P13_4_KINASE_3; 1.
SQ SEQUENCE 760 AA; 86888 MW; A081D76114080BE0 CRC64;

Query Match
Best Local Similarity 15.2%; Score 73; DB 5; Length 760;
Matches 24; Conservative 22; Mismatches 28; Indels 20; Gaps 5;

QY 4 QOWNFAGIEAASATQGN-----VTSIHLLDEGKQSL--TKLAAAGGSGSEAYQ 52
Db EQWH-EGLEEASRLYFGDRNVKGMFEILPLHMLERGQTLKTSFSQAYGRELTAYE 294

QY 53 GVQO-KWDATATELNN-----LQNLARTISE 78
Db MSQRYKTSVAVMDLDRWDIYHVFKISRQIPQ 328

RESULT 28
Q1555 PRELIMINARY; PRT; 1204 AA.
AC Q1555;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Probable bacteriophage protein.
GN PA0641.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=APCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.J., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR InterPro; IPR003961; FN_III.
DR ENBL; AE004499; AAG04030.1;
DR SMART; SM00060; FN3; 1.
KW Complete proteome.
SQ SEQUENCE 1204 AA; 130789 MW; 9090BE6D6712F759 CRC64;

Query Match
Best Local Similarity 15.2%; Score 73; DB 16; Length 1204;
Matches 23; Conservative 25; Mismatches 35; Indels 14; Gaps 4;

QY 11 IEAASATQGNVTGSIHLLDEGKQSL-----LTKLAAAGGSGS--EAYQGVQKWDATATE 64
Db LEATDAEGIAIEVATSDQATQRLDQLSASIGCTAASLQSFQTARANADSLAQ 1052

QY 65 LNNALQNLART-----ISEAGQMASTEGNVTMFA 95
Db RIDTVQ--ARTDTNSAATQTTTSQAVTSLDGNVAMYS 1087

RESULT 29
Q9VK45 PRELIMINARY; PRT; 2470 AA.
AC Q9VK45;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)

```

```

DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE CG5092 protein.
CN TOR OR CG5092.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Makos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Frannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Fabrics B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dedson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E., Wang X.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
DR ENBL; AE003638; AAF53237.1;
DR HSSP; P42345; 1FAP.
DR FlyBase; FBgn0021796; Tor.
DR InterPro; IPR003151; FAT.
DR InterPro; IPR003152; FATC.
DR InterPro; IPR000403; P13_P14_kinase.
DR InterPro; IPR001440; TPR.
DR Pfam; PF02259; FAT; 1.
DR Pfam; PF02260; FATC; 1.
DR Pfam; PF00454; P13_P14_kinase; 1.
DR SMART; SM00146; P13KC; 1.
DR PROSITE; PS00915; P13_4_KINASE_1; 1.
DR PROSITE; PS00916; P13_4_KINASE_2; 1.
DR PROSITE; PS00290; P13_4_KINASE_3; 1.
DR PROSITE; PS00290; P13_4_KINASE_3; 1.
SQ SEQUENCE 2470 AA; 281031 MW; 5D78D2ECC07C7FF9 CRC64;

Query Match
Best Local Similarity 15.2%; Score 73; DB 5; Length 2470;
Matches 24; Conservative 22; Mismatches 28; Indels 20; Gaps 5;

QY 4 QOWNFAGIEAASATQGN-----VTSIHLLDEGKQSL--TKLAAAGGSGSEAYQ 52
Db EQWH-EGLEEASRLYFGDRNVKGMFEILPLHMLERGQTLKTSFSQAYGRELTAYE 2004

```

QY 53 GYQQ-KWDATATLNN-----LQNLARTISE 78
 Db 2005 WSQRYKTSAVVMDLRANDIYHVFKISRQLPQ 2038

RESULT 30

O49013
 ID O49013 PRELIMINARY; PRT; 238 AA.
 AC O49013;
 DT 01-JUN-1998 (TREMELrel. 06, Created)
 DT 01-JUN-1998 (TREMELrel. 06, Last sequence update)
 DT 01-DEC-2001 (TREMELrel. 19, Last annotation update)
 DE Beta-1,3-glucanase 5 (Fragment).
 GN SGLU5.
 OS Glycine max (Soybean).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.
 OC NCB1_TaxID=3847;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Minsoy;
 RA MEDLINE=99403019; PubMed=10471725;
 RA Jin W., Horner H.T., Palmer R.G., Shoemaker R.C.;
 RT "Analysis and mapping of gene families encoding beta-1,3-glucanases of
 RT soybean."
 RL Genetics 153:445-452 (1999).
 DR EMEL; AF034110; AAC04712.1; -.
 DR HSPF; F15737; IGHS.
 DR InterPro; IPR000490; Glyco_hydro_17.
 DR Pfam; PF00332; Glyco_hydro_17; 1.
 FT NON_TER 1
 FT NON_TER 238
 SQ SEQUENCE 238 AA; 25537 MW; 66948CA8674968FB CRC64;

Query Match 15.1%; Score 72.5; DB 10; Length 238;
 Best Local Similarity 30.9%; Pred. No. 45;
 Matches 30; Conservative 15; Mismatches 35; Indels 17; Gaps 5;
 QY 12 EAAASATQG-NVTSIHSLDEGKSLTKLAAA--WGGSGSEAY-QGVQOKWDATATLNL- 66
 Db 38 EGVQLALRGSNIEVILGVDPNDQLSLTNAGATNWNKVKYKAYSONVKFYIAGNEIHP 97
 QY 67 -----NALQNLARTISEA---QAMASTEIGNVT 91
 Db 98 GDSLAGSVLPALENIQKAIKSAANLQGMKVSTAITDTT 134

Search completed: February 5, 2004, 17:40:02
 Job time : 17.7976 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:36:54 ; Search time 6.3333 Seconds
(without alignments)
634.663 Million cell updates/sec

Title: US-09-805-427A-1

Perfect score: 479

Sequence: 1 MTEQOWNFAGTEAAASATQ.....ISEAQWASIEGNTGMFA 95

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 306088

Minimum DB seq length: 6

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 90 summaries

Database : Issued Patents AA.*

- 1: /cgn2_6/ptodata/1/iaa/5A COMB.pap.*
- 2: /cgn2_6/ptodata/1/iaa/5B COMB.pap.*
- 3: /cgn2_6/ptodata/1/iaa/6A COMB.pap.*
- 4: /cgn2_6/ptodata/1/iaa/6B COMB.pap.*
- 5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pap.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	100.0	95	2	US-08-465-640-2
2	259	54.1	51	3	US-08-818-112-104
3	259	54.1	51	4	US-08-818-111-99
4	259	54.1	51	4	US-09-056-556-104
5	259	54.1	51	4	US-09-072-596-99
6	80	16.7	500	1	US-08-260-582-77
7	80	16.7	500	5	PCT-US95-05471-77
8	75	15.7	504	4	US-09-252-991A-26180
9	73	15.2	1207	4	US-09-252-991A-27470
10	71.5	14.9	94	4	US-09-073-009-21
11	71.5	14.9	94	4	US-09-073-009-31
12	71	14.8	94	4	US-09-073-009-23
13	70.5	14.7	94	4	US-09-073-009-19
14	70.5	14.7	94	4	US-09-073-009-29
15	70.5	14.7	331	4	US-09-223-040-4
16	68.5	14.3	94	4	US-09-073-009-27
17	67.5	14.1	475	4	US-09-252-991A-20838
18	67	14.0	445	4	US-09-477-135A-135
19	67	14.0	2293	3	US-09-368-590-2
20	66.5	13.9	94	4	US-09-073-009-25
21	66	13.8	185	2	US-08-691-814B-12
22	66	13.8	185	4	US-09-250-609-11
23	66	13.8	185	4	US-09-250-611-11
24	66	13.8	548	4	US-08-252-991A-23157
25	65.5	13.7	465	4	US-08-985-492-15
26	65.5	13.7	465	4	US-09-411-132A-6
27	65.5	13.7	525	4	US-09-187-999-9

ALIGNMENTS

RESULT 1

US-08-465-640-2
; Sequence 2, Application US/08465640
; Patent No. 5955077
; GENERAL INFORMATION:

28	65.5	13.7	537	4	US-09-252-991A-20929	Sequence 20929, A
29	65	13.6	95	4	US-08-311-731A-79	Sequence 79, Appl
30	65	13.6	112	4	US-08-311-731A-185	Sequence 185, App
31	65	13.6	319	4	US-09-198-452A-893	Sequence 893, App
32	65	13.6	1057	4	US-09-107-532A-4789	Sequence 4789, Ap
33	65	13.6	1151	4	US-09-328-352-4744	Sequence 4744, Ap
34	64.5	13.5	461	1	US-08-186-222-2	Sequence 2, Appli
35	64.5	13.5	476	2	US-08-955-713-4	Sequence 4, Appli
36	64.5	13.5	530	4	US-09-252-991A-23861	Sequence 23861, A
37	64.5	13.5	1043	4	US-08-851-567B-61	Sequence 61, Appl
38	64	13.4	189	1	US-08-884-682-4	Sequence 4, Appli
39	64	13.4	189	2	US-09-096-082-4	Sequence 4, Appli
40	64	13.4	293	4	US-09-252-991A-27425	Sequence 27425, A
41	64	13.4	562	3	US-09-012-515A-14	Sequence 14, Appl
42	64	13.4	562	3	US-08-360-144A-14	Sequence 14, Appl
43	64	13.4	562	4	US-09-012-504A-14	Sequence 14, Appl
44	64	13.4	562	4	US-09-012-399A-14	Sequence 14, Appl
45	64	13.4	562	4	PCT-US95-06722-14	Sequence 14, Appl
46	64	13.4	1209	4	US-09-252-991A-25844	Sequence 25844, A
47	64	13.4	3031	1	US-07-689-008-2	Sequence 2, Appli
48	63.5	13.3	690	4	US-09-252-991A-23187	Sequence 23187, A
49	63.5	13.3	975	4	US-09-914-259-19	Sequence 19, Appl
50	63.5	13.3	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
51	63	13.2	247	4	US-09-252-991A-29972	Sequence 29972, A
52	63	13.2	272	4	US-09-252-991A-28312	Sequence 28312, A
53	63	13.2	857	4	US-09-252-991A-23956	Sequence 23956, A
54	63	13.2	1477	1	US-08-038-682-4	Sequence 4, Appli
55	63	13.2	1477	1	US-08-302-832-4	Sequence 4, Appli
56	63	13.2	1477	2	US-08-530-198-4	Sequence 4, Appli
57	63	13.2	1477	2	US-08-469-880-4	Sequence 4, Appli
58	63	13.2	1477	2	US-08-728-470-4	Sequence 4, Appli
59	63	13.2	1477	2	US-08-617-697-4	Sequence 4, Appli
60	63	13.2	1477	3	US-08-719-641-4	Sequence 4, Appli
61	62.5	13.0	3729	2	US-08-804-227C-4	Sequence 4, Appli
62	62	12.9	130	4	US-09-732-210-1130	Sequence 1130, Ap
63	62	12.9	278	4	US-09-252-991A-21844	Sequence 21844, A
64	62	12.9	2504	4	US-09-328-352-5821	Sequence 5821, Ap
65	61.5	12.8	190	1	US-08-106-981-2	Sequence 2, Appli
66	61.5	12.8	316	4	US-09-308-375-3	Sequence 3, Appli
67	61.5	12.8	414	4	US-09-252-991A-31151	Sequence 31151, A
68	61.5	12.8	497	4	US-09-328-352-7959	Sequence 7959, Ap
69	61.5	12.8	1098	1	US-07-777-715-7	Sequence 7, Appli
70	61.5	12.8	1098	3	US-08-170-126-2	Sequence 2, Appli
71	61.5	12.8	1098	3	US-08-954-418-2	Sequence 2, Appli
72	61.5	12.8	1129	4	US-09-252-991A-29927	Sequence 2, Appli
73	61.5	12.8	4545	2	US-08-804-227C-14	Sequence 14, Appl
74	61.5	12.8	4550	2	US-08-804-227C-8	Sequence 8, Appli
75	61.5	12.8	4550	2	US-08-804-198-2	Sequence 2, Appli
76	61	12.7	304	3	US-09-105-390-40	Sequence 40, Appl
77	61	12.7	331	3	US-09-105-390-56	Sequence 56, Appl
78	61	12.7	347	4	US-09-252-991A-19006	Sequence 19006, A
79	61	12.7	375	3	US-09-863-743-1	Sequence 1, Appli
80	61	12.7	375	4	US-09-590-540-1	Sequence 1, Appli
81	61	12.7	418	3	US-09-141-000-6	Sequence 6, Appli
82	60.5	12.6	330	3	US-09-267-031-16	Sequence 16, Appl
83	60.5	12.6	338	1	US-08-891-254-1	Sequence 1, Appli
84	60.5	12.6	338	2	US-08-484-358-2	Sequence 2, Appli
85	60.5	12.6	338	2	US-08-819-539-1	Sequence 1, Appli
86	60.5	12.6	338	2	US-09-030-270A-1	Sequence 1, Appli
87	60.5	12.6	338	3	US-09-118-959-2	Sequence 2, Appli
88	60.5	12.6	338	3	US-08-984-207-1	Sequence 1, Appli
89	60.5	12.6	338	3	US-09-013-587-1	Sequence 1, Appli
90	60.5	12.6	338	4	US-09-086-118-21	Sequence 21, Appl

APPLICANT: ANDERSEN, Peter
APPLICANT: ANDERSEN, Ase Hengard
APPLICANT: HASLOV, Kaare
APPLICANT: SORESEN, Anne Lund
TITLE OF INVENTION: TUBERCULOSIS VACCINE
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Browdy and Neimark
STREET: 419 Seventh Street, N.W., Suite 300
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,640
FILING DATE: 05-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/123,182
FILING DATE: 20-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/DK94/00273
FILING DATE: 01-JUL-1994
ATTORNEY/AGENT INFORMATION:
NAME: COOPER, IVER P
REGISTRATION NUMBER: 28,005
REFERENCE/DOCKET NUMBER: ANDERSEN-3A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-3528
TELEFAX: 202-737-3528
TELEX: 248633
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 95 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-640-2

Query Match 100.0%; Score 479; DB 2; Length 95;
Best Local Similarity 100.0%; Pred. No. 6.3e-47;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 60
DB 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLLDEGKQSLTKLAAAWGGSGSEAYQGVQKQWDA 60
QY 61 TATELNALQNLARTISEAGQWASTEGNVTGMFA 95
DB 61 TATELNALQNLARTISEAGQWASTEGNVTGMFA 95

RESULT 2
US-08-818-112-104
Sequence 104, Application US/08818112
Patent No. 6290969
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillion, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR IMMUNOTHERAPY
AND DIAGNOSIS OF TUBERCULOSIS
NUMBER OF SEQUENCES: 153
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,112
FILING DATE: 13-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.411C6
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 104:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-818-112-104

Query Match 54.1%; Score 259; DB 3; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.9e-22;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLLDEGKQSLTKLAAAWGGSGSEAY 51
DB 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLLDEGKQSLTKLAAAWGGSGSEAY 51

RESULT 3
US-08-818-111-99
Sequence 99, Application US/08818111
Patent No. 633852
GENERAL INFORMATION:
APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillion, Davin C.
APPLICANT: Campos-Neto, Antonio
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
TUBERCULOSIS
NUMBER OF SEQUENCES: 148
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM: disk
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/818,111
FILING DATE: 13-MAR-1997
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C6
TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-818-111-99

Query Match 54.1%; Score 259; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.9e-22;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAWGGSGSEAY 51
|||||
Db 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAWGGSGSEAY 51

RESULT 4

US-09-056-556-104
Sequence 104, Application US/09056556
Patent No. 6350456

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE PREVENTION AND
NUMBER OF SEQUENCES: 241

CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/056,556
FILING DATE: 07-APR-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.457
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 104:

SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-056-556-104

Query Match 54.1%; Score 259; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.9e-22;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAWGGSGSEAY 51
|||||
Db 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAWGGSGSEAY 51

RESULT 5

US-09-072-596-99
Sequence 99, Application US/09072596
Patent No. 6458366

GENERAL INFORMATION:

APPLICANT: Reed, Steven G.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Dillon, Davin C.
APPLICANT: Campos-Neto, Antonia
APPLICANT: Houghton, Raymond
APPLICANT: Vedvick, Thomas S.
APPLICANT: Twardzik, Daniel R.
APPLICANT: Lodes, Michael J.
APPLICANT: Hendrickson, Ronald C.

TITLE OF INVENTION: COMPOUNDS AND METHODS FOR DIAGNOSIS OF
NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:

ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 99:

SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
US-09-072-596-99

Query Match 54.1%; Score 259; DB 4; Length 51;
Best Local Similarity 100.0%; Pred. No. 1.9e-22;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAWGGSGSEAY 51
|||||
Db 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAWGGSGSEAY 51

RESULT 6

US-08-260-582-77
Sequence 77, Application US/08260582
Patent No. 5635182

GENERAL INFORMATION:

APPLICANT: McCoy, John M.
APPLICANT: Lu, Zhijian
TITLE OF INVENTION: METHOD OF DETECTING LIGAND
INTERACTIONS
NUMBER OF SEQUENCES: 76
CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: Massachusetts
COUNTRY: U.S.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

RESULT 8
US-09-252-991A-26180
; Sequence 26180, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

RESULT 13
US-09-073-009-21
; Sequence 21, Application US/09073009
; Patent No. 655563
; GENERAL INFORMATION:
; APPLICANT: Alderson, Mark
; APPLICANT: Dillon, Davin C.
; APPLICANT: Skeiky, Yasir A.W.

APPLICANT: Campos-Neto, Antonio
TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Cumbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,009
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 21:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: linear
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-009-21

Query Match 14.9%; Score 71.5; DB 4; Length 94;
Best Local Similarity 25.5%; Pred. No. 0.64;
Matches 24; Conservative 20; Mismatches 31; Indels 19; Gaps 4;
QY 6 WNFAGIEAASAIQGNVTSHSLDEGKQSLTK--LAAA--WGGSGSEAYQG----- 53
Db 5 YQFGVDVHAHGAMIRAQA---GLLEABHQAIIRDVLVTASDFWGGAGSAACCGFTITQLGRN 60
QY 54 ---VQKWDATATLNNALQNLARTISEAGQAMA 84
Db 61 FQVIVEQANAHGQKVQAGNNAQTDSAVGSSWA 94

RESULT 11
US-09-073-009-31
Sequence 31, Application US/09073009
Patent No. 6555653
GENERAL INFORMATION:
APPLICANT: Alderson, Mark
APPLICANT: Dillon, Davin C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Campos-Neto, Antonio
TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Cumbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

TUBERCULOSIS AND ME

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,009
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-009-31

Query Match 14.9%; Score 71.5; DB 4; Length 94;
Best Local Similarity 25.5%; Pred. No. 0.64;
Matches 24; Conservative 20; Mismatches 31; Indels 19; Gaps 4;
QY 6 WNFAGIEAASAIQGNVTSHSLDEGKQSLTK--LAAA--WGGSGSEAYQG----- 53
Db 5 YQFGVDVHAHGAMIRAQA---GLLEABHQAIIRDVLVTASDFWGGAGSAACCGFTITQLGRN 60
QY 54 ---VQKWDATATLNNALQNLARTISEAGQAMA 84
Db 61 FQVIVEQANAHGQKVQAGNNAQTDSAVGSSWA 94

RESULT 12
US-09-073-009-23
Sequence 23, Application US/09073009
Patent No. 6555653
GENERAL INFORMATION:
APPLICANT: Alderson, Mark
APPLICANT: Dillon, Davin C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Campos-Neto, Antonio
TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Cumbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,009
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid

TUBERCULOSIS AND ME

TUBERCULOSIS AND ME

Fri Feb 6 13:54:43 2004

us-09-805-427a-1.116min.ra1

STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-009-23

Query Match 14.8%; Score 71; DB 4; Length 94;
Best Local Similarity 25.7%; Pred. No. 0.73;
Matches 26; Conservative 17; Mismatches 36; Indels 22; Gaps 5;

Qy 6 WNFAGIEAASAIQGNVTSHSLDEKQSLTK--LAAA--WGGSGSEAYQGVQKWDAT 61
Db 5 YQFGVDVDAHGMIRAQ--GLLEAEHQAIIRDVLTASDFWGGAGSACQGF----- 53
Qy 62 ATELNNALQNL-----ARTISEAGQAMASTEAGNVTCMFA 95
Db 54 ITQLGRNFQVIEQANTHGKQVQAAGNNMAQTDSAVXSSWA 94

RESULT 13
US-09-073-009-19
Sequence 19, Application US/09073009
Patent No. 6555653
GENERAL INFORMATION:
APPLICANT: Alderson, Mark
APPLICANT: Dillon, David C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Campos-Neto, Antonio
TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,009
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-009-19

TUBERCULOSIS AND ME

Query Match 14.7%; Score 70.5; DB 4; Length 94;
Best Local Similarity 25.5%; Pred. No. 0.83;
Matches 24; Conservative 20; Mismatches 31; Indels 19; Gaps 4;

Qy 6 WNFAGIEAASAIQGNVTSHSLDEKQSLTK--LAAA--WGGSGSEA----- 50
Db 5 YQFGVDVDAHGMIRAQAS---LEAEHQAIIRDVLAAGDFWGGAGSVACQEFITQLGRN 60

Qy 51 YQGVQKWDATATLNNALONLARTISEAGQAMA 84
Db 61 FQVIYEQANAHGKQVQAAGNNMAQTDSAVGSSWA 94

RESULT 14
US-09-073-009-29
Sequence 29, Application US/09073009
Patent No. 6555653
GENERAL INFORMATION:
APPLICANT: Alderson, Mark
APPLICANT: Dillon, David C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Campos-Neto, Antonio
TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY
STREET: 6300 Columbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,009
FILING DATE: 05-MAY-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-009-29

Query Match 14.7%; Score 70.5; DB 4; Length 94;
Best Local Similarity 25.5%; Pred. No. 0.83;
Matches 24; Conservative 21; Mismatches 30; Indels 19; Gaps 4;

Qy 6 WNFAGIEAASAIQGNVTSHSLDEKQSLTK--LAAA--WGGSGSEAYQG----- 53
Db 5 YQFGVDVDAHGMIR---ALAGLEAEHQAIISDVLTASDFWGGAGSACQGFITQLGRN 60

Qy 54 ---VQKWDATATLNNALONLARTISEAGQAMA 84
Db 61 FQVIYEQANAHGKQVQAAGNNMAQTDSAVGSSWA 94

RESULT 15
US-09-223-040-4
Sequence 4, Application US/09223040
Patent No. 6544522
GENERAL INFORMATION:
APPLICANT: Skeiky, Yasir
APPLICANT: Alderson, Mark
APPLICANT: Campos-Neto, Antonio
APPLICANT: Corixa Corporation
TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens

TUBERCULOSIS AND M

; TITLE OF INVENTION: and Their Uses
 ; FILE REFERENCE: 014058-003010US
 ; CURRENT APPLICATION NUMBER: US/09/223,040
 ; CURRENT FILING DATE: 1998-12-30
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 331
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence:tri-fusion
 US-09-223-040-4

Query Match
 Best Local Similarity 14.7%; Score 70.5; DB 4; Length 331;
 Matches 24; Conservative 20; Mismatches 31; Indels 19; Gaps 4;
 QY 6 WNFAGIEAASAIQGNVTSIHSLDEGKQSLTK--LAAA--WGGSGSEA-----50
 Db 242 YQFGVDVDAHGMIRAAAS---LEAEHQAIYRDVLAAGDFWGGAGSVACQEFITQLGRN 297
 QY 51 YQGVQCKWDATATLNNALQNLARTISEAQMA 84
 Db 298 FQVIEQANAHGQKVQAAGNNAQTDSAVGSSWA 331

RESULT 16
 US-09-073-009-27
 ; Sequence 27, Application US/09073009
 ; Patent No. 6555653
 ; GENERAL INFORMATION:
 ; APPLICANT: Alderson, Mark
 ; APPLICANT: Dillon, Davin C.
 ; APPLICANT: Skeiky, Yasir A.W.
 ; APPLICANT: Campos-Neto, Antonio
 ; TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
 ; NUMBER OF SEQUENCES: 144
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: SEED AND BERRY
 ; STREET: 6300 Cumbia Center, 701 Fifth Ave.
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: US
 ; ZIP: 98104
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/073,009
 ; FILING DATE: 05-MAY-1998
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Maki, David J.
 ; REGISTRATION NUMBER: 31,392
 ; REFERENCE/DOCKET NUMBER: 210121.441C1
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 206-622-4900
 ; TELEFAX: 206-682-6031
 ; INFORMATION FOR SEQ ID NO: 27:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 94 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; ORIGINAL SOURCE:
 ; ORGANISM: Mycobacterium tuberculosis
 US-09-073-009-27

Query Match
 14.3%; Score 68.5; DB 4; Length 94;

Best Local Similarity 25.5%; Pred. No. 1.4;
 Matches 24; Conservative 19; Mismatches 32; Indels 19; Gaps 4;
 QY 6 WNFAGIEAASAIQGNVTSIHSLDEGKQSLTK--LAAA--WGGSGSEAYQG-----53
 Db 5 YQFGVDVDAHGMIRAXA---GLLEAEHQALISDLTASDFWGGAGSAACQGFITQLGRN 60
 QY 54 ---VQCKWDATATLNNALQNLARTISEAQMA 84
 Db 61 FQVIEQANAHGQKVQAAGNNAQTDSAVGSSWA 94

RESULT 17
 US-09-252-991A-20838
 ; Sequence 20838, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 20838
 ; LENGTH: 475
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-20838

Query Match
 Best Local Similarity 14.1%; Score 67.5; DB 4; Length 475;
 Matches 24; Conservative 17; Mismatches 38; Indels 11; Gaps 4;
 QY 6 WNFAGIEAASAIQGNVTSIHSLDEGKQSLTKLAAWGGSGSEAYQGVQCKWDATATEL 65
 Db 267 WHPAGVEAGGEGVGGNPRG-----DFAHPSQRRQAGA-GGPGAEQRQRQAHHGQPDQL 320
 QY 66 NNALQN--LARTISEAQMASTEAGNVTGM 93
 Db 321 AHALQEVLMVDVEDHGESRSV---VAGL 347

RESULT 18
 US-09-477-135A-135
 ; Sequence 135, Application US/09477135A
 ; Patent No. 6572865
 ; GENERAL INFORMATION:
 ; APPLICANT: Nano, Francis
 ; TITLE OF INVENTION: Mycobacterium Tuberculosis DNA Sequences Encoding
 ; TITLE OF INVENTION: Immunostimulatory Peptides
 ; FILE REFERENCE: 52888
 ; CURRENT APPLICATION NUMBER: US/09/477,135A
 ; CURRENT FILING DATE: 2000-01-03
 ; PRIOR APPLICATION NUMBER: 08990823
 ; PRIOR FILING DATE: 1997-12-15
 ; PRIOR APPLICATION NUMBER: US 96/10375
 ; PRIOR FILING DATE: 1996-06-14
 ; PRIOR APPLICATION NUMBER: 60/000,254
 ; NUMBER OF SEQ ID NOS: 169
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 135
 ; LENGTH: 445
 ; TYPE: PRT
 ; ORGANISM: Mycobacterium tuberculosis
 US-09-477-135A-135

Query Match
 14.0%; Score 67; DB 4; Length 445;

TUBERCULOSIS AND ME

Fri Feb 6 13:54:43 2004

ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.441C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 206-622-4900
TELEFAX: 206-682-6031
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 94 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
ORIGINAL SOURCE:
ORGANISM: Mycobacterium tuberculosis
US-09-073-009-25

Query Match 13.9%; Score 66.5; DB 4; Length 94;
Best Local Similarity 23.7%; Pred. No. 2.3;
Matches 22; Conservative 20; Mismatches 34; Indels 17; Gaps 3;

QY 6 WNFAGIEAASAIQGNVTSIHSLLDEGKQSLTKLAAA---WGGSGSEAYQG----- 53
Db 5 YQFGVDVDAHGAMIRAQAGSLEA---EHOAISDVLTASDFWGGAGSAAACQGFITQLGRNF 61
QY 54 --VQOKWDATATLNNALQNLARTISEAGQAWA 84
Db 62 QVXYEQANAHGQKVOAGNNAQTDSAVGSSWA 94

RESULT 21
US-08-691-814B-12
Sequence 12, Application US/08691814B
Patent No. 5981218
GENERAL INFORMATION:
APPLICANT: Rio, Marie-Christine
APPLICANT: Tomasetto, Catherine
APPLICANT: Bassett, Paul
APPLICANT: Byrne, Jennifer
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Useful
NUMBER OF SEQUENCES: 124
CORRESPONDENCE ADDRESS:
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
STREET: 1100 New York Ave, NW, Suite 600
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/691,814B
FILING DATE: 31-JUL-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/002,183
FILING DATE: 09-AUG-1995
ATTORNEY/AGENT INFORMATION:
NAME: Steffe, Eric K.
REGISTRATION NUMBER: 36,688
REFERENCE/DOCKET NUMBER: 1383.0090001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-371-2600
TELEFAX: 202-371-2543
INFORMATION FOR SEQ ID NO: 12:
SEQUENCE CHARACTERISTICS:
LENGTH: 185 amino acids
TYPE: amino acid

Best Local Similarity 25.9%; Pred. No. 16;
Matches 22; Conservative 14; Mismatches 33; Indels 16; Gaps 2;

QY 9 AGIEAASAIQGNVTSIHSLLDEGKQSLTKLAAWGGSGSEAYQGVQKWDATATLNNNA 68
Db 159 AAINQAADTLDGNGDSLHNLRE---LAQVAGRIQDSRGDIFGTGKN----- 202

QY 69 LONLARTISEAGQAWASTEGNVUTQM 93
Db 203 LQVLVDALSESDEQIVQFAGHVASV 227

RESULT 19
US-09-368-590-2
Sequence 2, Application US/09368590
Patent No. 6187563
GENERAL INFORMATION:
APPLICANT: Solimena, Michele
TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
FILE REFERENCE: AUTOANTIGENS OF AUTOIMMUNE DISEASES
FILE REFERENCE: 101918-200 (OCR-941)
CURRENT APPLICATION NUMBER: US/09/368,590
CURRENT FILING DATE: 1999-08-04
EARLIER FILING DATE: 1998-08-07
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 2293
TYPE: PRT
ORGANISM: Human
US-09-368-590-2

Query Match 14.0%; Score 67; DB 3; Length 2293;
Best Local Similarity 26.2%; Pred. No. 1.4e+02;
Matches 21; Conservative 16; Mismatches 29; Indels 14; Gaps 4;

QY 5 QWNFAGIEAASAIQGNVTSIHSLLDEGKQSLTKLAAWGGSGSEAYQGVQ---KWDAT 61
Db 719 QWRLSGLRALQALEPQRA---ALLEEA---ALLAERFPACAAXLHOGAEIIGAEMGAL 771
QY 62 ATELNALQNLARTISEAGQ 81
Db 772 A-----SAAQACGEAAVAAAGR 787

RESULT 20
US-09-073-009-25
Sequence 25, Application US/09073009
Patent No. 6555653
GENERAL INFORMATION:
APPLICANT: Alderson, Mark
APPLICANT: Dillon, David C.
APPLICANT: Skeiky, Yasir A.W.
APPLICANT: Campos-Neto, Antonio
TITLE OF INVENTION: COMPOUNDS AND DIAGNOSIS OF
NUMBER OF SEQUENCES: 144
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY
STREET: 6300 Cumbia Center, 701 Fifth Ave.
CITY: Seattle
STATE: Washington
COUNTRY: US
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/073,009
FILING DATE: 05-MAY-1998
CLASSIFICATION:

TUBERCULOSIS AND ME


```

; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-691-814B-12

```

Query Match 13.8%; Score 66; DB 2; Length 185;
Best Local Similarity 28.3%; Pred. No. 6.5;
Matches 22; Conservative 19; Mismatches 29; Indels

Qy	18	I QGNVTSIHSLDEKQSILTKLAAMWGGSEAY-QGVQQKW-DATATELNNAQLNART	75
Dd	46	VVEEIQTLISVLAAKEKHIAELKRKLGISSIQEFKNIAKGWDVTAT---	102
Qy	76	ISEAGQ---AMASTEENV	90
Dd	103	LSQAGOKASAAFSSVGSV	120

RESULT 22

```

US-09-250-609-11
; Sequence 11, Application US/09250609A
; Patent No. 6458943
; GENERAL INFORMATION:
; APPLICANT: Byrne, Jennifer A.
; TITLE OF INVENTION: Members of the D
; FILE REFERENCE: 1383.0210002
; CURRENT APPLICATION NUMBER: US/09/250
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-250-609-11

```

	Query Match	13.8%	Score 66	DB 4	Length 185
	Best Local Similarity	28.2%	Pred. No. 6.5		
	Matches	22	Conservative	19	Mismatches 29
					Indels 8
					Gaps 4
Qy	18	IQGVNTSIHSLDEGKOSLTKLAAAGGSGSEAY-QGVQOKW-DATATELNNAQNIART	75		
Db	46	VEEIEITQSVLAAKEHKLAEKRLKGISSLQFKQNIAGQWDVAT--NAVYKTSSET	102		

```
QY      76 ISEAGQ---AMASTE GNV 90
        :|:|:| | | | | | | |
Db     103 L SQAGQKASAA FSSVGSV 120
```

RESULT 23

```

US-09-250-611-11
; Sequence 11, Application US/09250611
; Patent No. 6528283
; GENERAL INFORMATION:
; APPLICANT: Byrnes, Jennifer A.
; APPLICANT: Basset, Paul
; TITLE OF INVENTION: Members of the D52 Gene Family
; FILE REFERENCE: 1383.0210001
; CURRENT APPLICATION NUMBER: US/09/250,611
; CURRENT FILING DATE: 1999-02-17
; NUMBER OF SEQ ID NOS: 108
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 11
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-250-611-11

```

```
Query Match      13.8%; Score 66; DB 4; Length 185;
Best Local Similarity 28.2%; Pred. No. 6.5;
Matches 22; Conservative 19; Mismatches 29; Indels 8; Gaps 4;

18 ICGNVTSIHSLDEGKQSLTKLAANGGSGSEAY-QGVQOKW-DATATELNNAQLNART 75
:: : : : : : | : : : : : | : : : : : | : : : : :
```

```

Db      46  VEETIOTLSVLAKEKHIAELKRLGISSLOEFKQNTAKGWQDVAT--NAYKKTSET 102

QY      76  ISEAGQ--AMASTEQNV 90
        :|:|:| | | :|:|
Db      103 LSAQAGQKASAAAFSSVGSV 120

```

RESULT 24

```

US-09-252-991A-23157
; Sequence 23157, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AN
; TITLE OF INVENTION: AERUGINOSA FOR DIA
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23157
; LENGTH: 548
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23157

```

Query Match	13.8%	Score 66;	DB 4;	Length 548;
Best Local Similarity	27.9%;	Pred. No. 27;		
Matches	31;	Conservative 11;	Mismatches 47;	Indels 22; Gaps 5;

QY	2	TEQWNFAG-----	TEAASAIQGNVTS	IHSLDEGKQSLTK	LAAGWGSGBAYQGVQ	56
Db	18	TRAYW-YAGSAGSPV	AGLSLTGTACRSF	PEDEWNSQANELL	IGAGPDG-----	QVVG 72
QY	57	KW-----	DATATELNNALON	LARTISAGQAM--	ASTEGNVTGMFA	95
Db	73	ANRLANRHGLI	CAGTGKTVTLQHL	AESFSDAGVAFAD	IKDLCGIAA	123

RESULT 25

```

US-08-385-492-15
Sequence 15, Application US/08985492
Patent No. 6395530
GENERAL INFORMATION:
APPLICANT: Jaye, Michael C.
APPLICANT: Doan, Kim-Anh T.
APPLICANT: Krawiec, John A.
APPLICANT: Lynch, Kevin J.
APPLICANT: Amin, Dilip V.
APPLICANT: South, Victoria J.
TITLE OF INVENTION: LLG POLYPEPTIDES
TITLE OF INVENTION: LIPASE FAMILY, AN
TITLE OF INVENTION: IN ENZYMATIC HYDR
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESS: Rhone-Poulenc Rorer Inc.
STREET: 500 Arcola Rd. 3C43
CITY: Collegeville
STATE: PA
COUNTRY: USA
ZIP: 19426
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Ve
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/985,492
FILING DATE:
CLASSIFICATION: 800

```

ATTORNEY/AGENT INFORMATION:
 NAME: Fehner Ph.D., Paul F.
 REGISTRATION NUMBER: 35,135
 REFERENCE/DOCKET NUMBER: A2582-US
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (610)454-3839
 TELEFAX: (610)454-3808
 INFORMATION FOR SEQ ID NO: 15:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 465 amino acids
 TYPE: amino acid
 STRANDEDNESS:
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-985-492-15

Query Match 13.7%; Score 65.5; DB 4; Length 465;
 Best Local Similarity 19.5%; Pred. No. 25;
 Matches 25; Conservative 16; Mismatches 46; Indels 41; Gaps 3;
 QY 7 NFAGIEAASAIQGNVTS-----IHSLLDEGK-----QSLTKLAAAWG 44
 Db 65 NFQEAADSSISGSGNFKNRKTFRILHGFIDKGEENLANVCKNLFKVESVNCICVDWK 124
 QY 45 GSGSEAYQGVQKWDATATLNNALQNL-----ARTISEAGQAWAS 85
 Db 125 GGSRTGYTQASQNIIRIVGAEVYFVEFLOSAFGYSPSNVHVHIGSLGHAHAGEAGRTNG 184
 QY 86 TEGNVITGM 93
 Db 185 TIGRTIGL 192

RESULT 26
 US-09-411-132A-6
 ; Sequence 6, Application US/09411132A
 ; Patent No. 6558936
 ; GENERAL INFORMATION:
 ; APPLICANT: Rhodadoust, Mehran
 ; APPLICANT: Kapeller-Libermann, Rosana
 ; TITLE OF INVENTION: No. 6558936el Human Lipase Proteins, Nucleic Acides Encoding
 ; TITLE OF INVENTION: Them, and Uses of Both of These
 ; FILE REFERENCE: 10147-14
 ; CURRENT APPLICATION NUMBER: US/09/411,132A
 ; PRIOR FILING DATE: 2000-09-12
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 6
 ; LENGTH: 465
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-411-132A-6

Query Match 13.7%; Score 65.5; DB 4; Length 465;
 Best Local Similarity 19.5%; Pred. No. 25;
 Matches 25; Conservative 16; Mismatches 46; Indels 41; Gaps 3;
 QY 7 NFAGIEAASAIQGNVTS-----IHSLLDEGK-----QSLTKLAAAWG 44
 Db 65 NFQEAADSSISGSGNFKNRKTFRILHGFIDKGEENLANVCKNLFKVESVNCICVDWK 124
 QY 45 GSGSEAYQGVQKWDATATLNNALQNL-----ARTISEAGQAWAS 85
 Db 125 GGSRTGYTQASQNIIRIVGAEVYFVEFLOSAFGYSPSNVHVHIGSLGHAHAGEAGRTNG 184
 QY 86 TEGNVITGM 93
 Db 185 TIGRTIGL 192

RESULT 27
 US-09-187-999-9
 ; Sequence 9, Application US/09187999A

Patent No. 6482646
 GENERAL INFORMATION:
 APPLICANT: Gindullis, Frank
 APPLICANT: Meier, Iris
 TITLE OF INVENTION: Plant Proteins that Interact with Nuclear Matrix
 TITLE OF INVENTION: Proteins and Function as Transcriptional Activators
 FILE REFERENCE: CL-1321
 CURRENT APPLICATION NUMBER: US/09/187,999A
 CURRENT FILING DATE: 1998-11-06
 NUMBER OF SEQ ID NOS: 39
 NUMBER OF SEQ ID NOS: 39
 SOFTWARE: Microsoft Office 97
 SEQ ID NO 9
 LENGTH: 525
 TYPE: PRT
 ORGANISM: Lycopersicon esculentum
 US-09-187-999-9

Query Match 13.7%; Score 65.5; DB 4; Length 525;
 Best Local Similarity 29.1%; Pred. No. 29;
 Matches 34; Conservative 13; Mismatches 41; Indels 29; Gaps 6;
 QY 1 MTEQQWNT-AGIEAASAIQ-----GNVTSI-HSLLEDEG-----KQSLTKLAA 41
 Db 275 LTESIFNTHRAAGEALASIQKEIGKDTNFGHSLVDGIPSNFTSDSDDFRSSTSVL- 333
 QY 42 AWGSGSGSEAYQGVQKWDATATLNNALQN-----LARTISEAGQAWASTEGNVITGM 95
 Db 334 -----SQNRDGAQSSLSNSEVYSLKQNIILLQNVQEAADLAKSEARVTELEA 384

RESULT 28
 US-09-252-991A-20929
 ; Sequence 20929, Application US/09252991A
 ; Patent No. 6551795
 ; GENERAL INFORMATION:
 ; APPLICANT: Marc J. Rubenfield et al.
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
 ; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 107196.136
 ; CURRENT APPLICATION NUMBER: US/09/252,991A
 ; CURRENT FILING DATE: 1999-02-18
 ; PRIOR APPLICATION NUMBER: US 60/074,788
 ; PRIOR FILING DATE: 1998-02-18
 ; PRIOR APPLICATION NUMBER: US 60/094,190
 ; PRIOR FILING DATE: 1998-07-27
 ; NUMBER OF SEQ ID NOS: 33142
 ; SEQ ID NO 20929
 ; LENGTH: 537
 ; TYPE: PRT
 ; ORGANISM: Pseudomonas aeruginosa
 US-09-252-991A-20929

Query Match 13.7%; Score 65.5; DB 4; Length 537;
 Best Local Similarity 21.7%; Pred. No. 30;
 Matches 20; Conservative 16; Mismatches 47; Indels 9; Gaps 1;
 QY 11 IEAASAIQGNVTSIHSLLDEGKQSLTKLAAAWGSGSEAYQGVQ-----KWDAT 61
 Db 409 VRSIANPTQOSTQEIQGLIEQLQDQANDAVAAAMRGSAHQSNLVEADSAQAALGRIVAT 468
 QY 62 ATLNNALQNARTISEAGQAWASTEGNVITGM 93
 Db 469 VEELDGLNQIATAAEQSQVAQDIDRNITNV 500

RESULT 29
 US-08-311-731A-79
 ; Sequence 79, Application US/08311731A
 ; Patent No. 6583266
 ; GENERAL INFORMATION:
 ; APPLICANT: SMITH, DOUGLAS
 ; APPLICANT: MAO, JEN-I
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES

RESULT 30
 US-08-311-731A-185
 ; Sequence 185, Application US/08311731A
 ; Patent No. 6583266
 ; GENERAL INFORMATION:
 ; APPLICANT: SMITH, DOUGLAS
 ; APPLICANT: MAO, JEN-I
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
 ; TITLE OF INVENTION: RELATING TO MYCOBACTERIUM TUBERCULOSIS AND LAPRAE FOR
 ; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
 ; NUMBER OF SEQUENCES: 411
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: WOLF, GREENFIELD & SACKS, P.C.
 ; STREET: 600 ATLANTIC AVENUE
 ; CITY: BOSTON
 ; STATE: MASSACHUSETTS
 ; COUNTRY: USA
 ; ZIP: 02210
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25

Search completed: February 5, 2004, 17:41:21
Job time : 7.33333 secs

```

RESULT 30
US-08-311-731A-185
; Sequence 185, Application US/08311731A
; Patent No. 6583266
;
; GENERAL INFORMATION:
;
; APPLICANT: SMITH, DOUGLAS
;
; APPLICANT: MAO, JEN-I
;
; TITLE OF INVENTION: NUCLEIC ACID AN
;
; TITLE OF INVENTION: RELATING TO MTC
;
; TITLE OF INVENTION: DIAGNOSTICS AND
;
; NUMBER OF SEQUENCES: 411
;
; CORRESPONDENCE ADDRESS:
;
; ADDRESSEE: WOLF, GREENFIELD & SAC
;
; STREET: 600 ATLANTIC AVENUE
;
; CITY: BOSTON
;
; STATE: MASSACHUSETTS
;
; COUNTRY: USA
;
; ZIP: 02210
;
; COMPUTER READABLE FORM:
;
; MEDIUM TYPE: Floppy disk
;
; COMPUTER: IBM PC compatible
;
; OPERATING SYSTEM: PC-DOS/MS-DOS
;
; SOFTWARE: PatentIn Release #1.0

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 5, 2004, 17:38:58 ; Search time 13.119 Seconds
(without alignments)
1516.217 Million cell updates/sec

Title: US-09-805-427A-1
Perfect score: 479
Sequence: 1 MTEQWNFAGIEAASAIQ.....ISERAGMASTEGNVTGMFA 95

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 801455 seqs, 209382283 residues

Total number of hits satisfying chosen parameters: 788294

Minimum DB seq length: 6
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 90 summaries

Database :

Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pap.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pap.*
3: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pap.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pap.*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pap.*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pap.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pap.*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pap.*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap.*
12: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap.*
13: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pap.*
14: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pap.*
15: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pap.*
16: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap.*
17: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pap.*
18: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pap.*
19: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	479	100.0	95	10	US-09-805-427A-1
2	479	100.0	95	12	US-09-872-505-1
3	479	100.0	403	9	US-09-791-171-173
4	479	100.0	403	10	US-09-805-427A-4
5	479	100.0	403	12	US-09-804-980-173
6	479	100.0	403	12	US-09-872-505-4
7	479	100.0	404	9	US-09-791-171-172
8	479	100.0	404	10	US-09-805-427A-3
9	479	100.0	404	12	US-09-804-980-172
10	479	100.0	404	12	US-09-872-505-3
11	259	54.1	51	12	US-10-084-843-104
12	259	54.1	51	12	US-10-193-002-99
13	259	54.1	51	12	US-10-098-732A-33
14	186	38.8	95	16	US-10-080-170-11
15	151	31.5	96	10	US-09-738-626-4145

16	101	21.1	20	10	US-09-813-333-60	Sequence 60, Appl
17	101	21.1	20	14	US-10-044-703-60	Sequence 60, Appl
18	98	20.5	20	10	US-09-813-333-62	Sequence 62, Appl
19	98	20.5	20	14	US-10-044-703-62	Sequence 62, Appl
20	97	20.3	20	10	US-09-813-333-61	Sequence 61, Appl
21	97	20.3	20	14	US-10-044-703-61	Sequence 61, Appl
22	96	20.0	20	10	US-09-804-980-175	Sequence 175, Appl
23	96	20.0	20	12	US-09-804-980-199	Sequence 199, Appl
24	89	18.6	99	15	US-10-156-761-10069	Sequence 10069, A
25	85	17.7	96	9	US-09-791-171-2	Sequence 2, Appl
26	85	17.7	96	12	US-09-804-980-2	Sequence 2, Appl
27	85	17.7	96	12	US-09-804-980-195	Sequence 195, Appl
28	85	17.7	96	16	US-10-080-170-378	Sequence 378, Appl
29	81.5	17.0	108	10	US-09-738-626-4144	Sequence 4144, Appl
30	81	16.9	15	10	US-09-916-201-7	Sequence 7, Appl
31	81	16.9	96	16	US-10-080-170-299	Sequence 299, Appl
32	79	16.5	15	12	US-10-079-167-79	Sequence 79, Appl
33	79	16.5	15	12	US-10-345-000-1	Sequence 1, Appl
34	78.5	16.4	368	9	US-09-287-849-24	Sequence 24, Appl
35	78.5	16.4	368	12	US-10-359-460-24	Sequence 24, Appl
36	78.5	16.4	600	9	US-09-387-849-22	Sequence 22, Appl
37	78.5	16.4	600	12	US-10-359-460-22	Sequence 22, Appl
38	77	16.1	15	12	US-10-079-167-80	Sequence 80, Appl
39	77	16.1	15	12	US-10-345-000-2	Sequence 2, Appl
40	76.5	16.0	837	9	US-09-815-242-5883	Sequence 5883, Appl
41	76.5	16.0	875	9	US-09-815-242-13080	Sequence 13080, A
42	76.5	16.0	2434	9	US-09-815-242-5835	Sequence 5835, Appl
43	76.5	16.0	6281	9	US-09-815-242-12996	Sequence 12996, A
44	76	15.9	16	10	US-09-813-333-63	Sequence 63, Appl
45	76	15.9	16	14	US-10-044-703-63	Sequence 63, Appl
46	74.5	15.6	364	10	US-09-756-983-22	Sequence 22, Appl
47	74	15.4	644	12	US-10-369-493-11283	Sequence 11283, A
48	73	15.2	2086	9	US-09-815-242-5629	Sequence 5629, Appl
49	73	15.2	5795	9	US-09-815-242-12610	Sequence 12610, A
50	72.5	15.1	239	12	US-09-287-849-19	Sequence 19, Appl
51	72.5	15.1	239	9	US-10-359-460-19	Sequence 19, Appl
52	72.5	15.1	239	12	US-10-098-732A-47	Sequence 47, Appl
53	72.5	15.1	433	9	US-09-287-849-14	Sequence 14, Appl
54	72.5	15.1	433	12	US-10-359-460-14	Sequence 14, Appl
55	72.5	15.1	710	9	US-09-287-849-16	Sequence 16, Appl
56	72.5	15.1	710	12	US-10-359-460-16	Sequence 16, Appl
57	72.5	15.1	710	12	US-10-098-732A-49	Sequence 49, Appl
58	72.5	15.1	856	9	US-09-287-849-12	Sequence 12, Appl
59	72.5	15.1	856	12	US-10-359-460-12	Sequence 12, Appl
60	72	15.0	15	10	US-09-916-201-5	Sequence 5, Appl
61	72	15.0	573	15	US-10-043-487-327	Sequence 327, Appl
62	72	15.0	685	12	US-10-158-034-96	Sequence 96, Appl
63	72	15.0	1225	12	US-10-158-034-76	Sequence 76, Appl
64	71.5	14.9	66	9	US-09-864-761-33985	Sequence 33985, A
65	71.5	14.9	94	9	US-09-073-009-21	Sequence 21, Appl
66	71.5	14.9	94	9	US-09-073-009-31	Sequence 31, Appl
67	71.5	14.9	94	9	US-09-023-588-21	Sequence 21, Appl
68	71.5	14.9	94	9	US-09-023-588-31	Sequence 31, Appl
69	71.5	14.9	94	9	US-09-793-306-21	Sequence 21, Appl
70	71.5	14.9	94	9	US-09-793-306-31	Sequence 31, Appl
71	71	14.8	94	9	US-09-073-009-23	Sequence 23, Appl
72	71	14.8	94	9	US-09-023-588-23	Sequence 23, Appl
73	71	14.8	94	9	US-09-793-306-23	Sequence 23, Appl
74	70.5	14.7	94	9	US-09-073-009-19	Sequence 19, Appl
75	70.5	14.7	94	9	US-09-073-009-29	Sequence 29, Appl
76	70.5	14.7	94	9	US-09-023-588-19	Sequence 19, Appl
77	70.5	14.7	94	9	US-09-023-588-29	Sequence 29, Appl
78	70.5	14.7	94	9	US-09-793-306-19	Sequence 19, Appl
79	70.5	14.7	94	9	US-09-793-306-29	Sequence 29, Appl
80	70.5	14.7	94	12	US-10-098-732A-27	Sequence 27, Appl
81	70.5	14.7	94	16	US-10-080-170-611	Sequence 611, Appl
82	70.5	14.7	331	12	US-09-287-849-4	Sequence 4, Appl
83	70.5	14.7	331	12	US-10-359-460-4	Sequence 4, Appl
84	70.5	14.7	331	12	US-10-369-983-4	Sequence 4, Appl
85	70.5	14.7	825	12	US-10-369-983-14	Sequence 14, Appl
86	70	14.6	265	10	US-09-738-626-5873	Sequence 5873, Appl
87	69.5	14.5	1029	9	US-09-815-242-5885	Sequence 5885, Appl
88	69.5	14.5	1048	9	US-09-815-242-13083	Sequence 13083, A

89 69 14.4 15 10 US-09-916-201-6 Sequence 6, Appli
90 69 14.4 221 10 US-09-738-626-6507 Sequence 6507, Ap

ALIGNMENTS

```
RESULT 1
US-09-805-427A-1
; Sequence 1, Application US/09805427A
; Patent No. US20020176867A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Hybrids of M. tuberculosis Antigens
; FILE REFERENCE: 670001-2002.5
; CURRENT APPLICATION NUMBER: US/09/805,427A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-805-427A-1
Query Match      100.0%; Score 479; DB 10; Length 95;
Best Local Similarity 100.0%; Pred. No. 9.2e-45;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQOKWDA 60
Db 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQOKWDA 60

Qy 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95
Db 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95

RESULT 2
US-09-872-505-1
; Sequence 1, Application US/09872505
; Publication No. US20040013685A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Nucleic Acid Fragments Derived From M. Tuberculosis
; FILE REFERENCE: 670001-2002.6
; CURRENT APPLICATION NUMBER: US/09/872,505
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 95
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-872-505-1
Query Match      100.0%; Score 479; DB 12; Length 95;
Best Local Similarity 100.0%; Pred. No. 9.2e-45;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQOKWDA 60
Db 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQOKWDA 60

Qy 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95
Db 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95

RESULT 3
US-09-791-171-173
; Sequence 173, Application US/09791171
; Patent No. US20020094336A1
```

```
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDRINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 173
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-173
Query Match      100.0%; Score 479; DB 9; Length 403;
Best Local Similarity 100.0%; Pred. No. 5.6e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQOKWDA 60
Db 22 MTEQOWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQOKWDA 81

Qy 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95
Db 82 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 116

RESULT 4
US-09-805-427A-4
; Sequence 4, Application US/09805427A
; Patent No. US20020176867A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Hybrids of M. tuberculosis Antigens
; FILE REFERENCE: 670001-2002.5
; CURRENT APPLICATION NUMBER: US/09/805,427A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant Fusion protein ESAT-6-Ags85B
US-09-805-427A-4
Query Match      100.0%; Score 479; DB 10; Length 403;
Best Local Similarity 100.0%; Pred. No. 5.6e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQOKWDA 60
Db 22 MTEQOWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQOKWDA 81

Qy 61 TATELNALQNLARTISEAGQAMASTEAGNVTGMFA 95
```

```

Db      82 TATELNALQNLARTISEAGQAMASTEGNVTGMFA 116
|||||
RESULT 5
US-09-804-980-173
; Sequence 173, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Andersen, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 173
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-173

Query Match      100.0%; Score 479; DB 12; Length 403;
Best Local Similarity 100.0%; Pred. No. 5.6e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MTEQQWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQKQWDA 60
Db      22 MTEQQWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQKQWDA 81
QY      61 TATELNALQNLARTISEAGQAMASTEGNVTGMFA 95
Db      82 TATELNALQNLARTISEAGQAMASTEGNVTGMFA 116

RESULT 6
US-09-872-505-4
; Sequence 4, Application US/09872505
; Publication No. US20040013685A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Nucleic Acid Fragments Derived From M. Tuberculosis
; FILE REFERENCE: 670001-2002.6
; CURRENT APPLICATION NUMBER: US/09/872,505
; CURRENT FILING DATE: 2001-06-01
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant Fusion protein ESAT-6-Ag85B
US-09-872-505-4

Query Match      100.0%; Score 479; DB 12; Length 403;
Best Local Similarity 100.0%; Pred. No. 5.6e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MTEQQWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQKQWDA 60
Db      22 MTEQQWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQKQWDA 81
QY      61 TATELNALQNLARTISEAGQAMASTEGNVTGMFA 95
Db      82 TATELNALQNLARTISEAGQAMASTEGNVTGMFA 116

RESULT 7
US-09-791-171-172
; Sequence 172, Application US/09791171
; Patent No. US20020094336A1

```

```

; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKRANDS, Ida
; APPLICANT: WELDLING, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 172
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-172

```

```

Query Match      100.0%; Score 479; DB 2; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.6e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MTEQQWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQKQWDA 60
Db      310 MTEQQWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQKQWDA 369
QY      61 TATELNALQNLARTISEAGQAMASTEGNVTGMFA 95
Db      370 TATELNALQNLARTISEAGQAMASTEGNVTGMFA 404

```

```

RESULT 8
US-09-805-427A-3
; Sequence 3, Application US/09805427A
; Patent No. US20020176867A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; TITLE OF INVENTION: Hybrids of M. tuberculosis Antigens
; FILE REFERENCE: 670001-2002.5
; CURRENT APPLICATION NUMBER: US/09/805,427A
; CURRENT FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3
; LENGTH: 404
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Recombinant Fusion protein Ag85B-ESAT-6
US-09-805-427A-3

```

```

Query Match      100.0%; Score 479; DB 10; Length 404;
Best Local Similarity 100.0%; Pred. No. 5.6e-44;
Matches 95; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MTEQQWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQKQWDA 60
Db      310 MTEQQWNFAGIEAAASAIQGNVTSIHSLDDEGKQSLTKLAAAWGSGSEAYQGVQKQWDA 369
QY      61 TATELNALQNLARTISEAGQAMASTEGNVTGMFA 95

```

RESULT 11
US-10-084-843-104
; Sequence 104, Application US/10084843
: Publication No. US20030143243A1

TITLE OF INVENTION: COMBINATION OF
TUBERCULOSIS

NUMBER OF SEQUENCES: 350
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
City: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/193,002
FILING DATE: 10-Jul-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/072,596
FILING DATE: 05-MAY-1998
ATTORNEY/AGENT INFORMATION:
NAME: Maki, David J.
REGISTRATION NUMBER: 31,392
REFERENCE/DOCKET NUMBER: 210121.417C9
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 99:
SEQUENCE CHARACTERISTICS:
LENGTH: 51 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 99:
US-10-193-002-99

```

Query Match      54.1%; Score 259; DB 12; Length 51;
Best Local Similarity 100.0%; Pred. No. 4.2e-21;
Matches 51; Conservative 0; Mismatches 0; Indels 0;

Qy 1 MTEQWNFGAIEAAASAIQGNVTSIISLIDEGKQSITKLAAGWGSGSSRAY 51
Db 1 MTEQWNFGAIEAAASAIQGNVTSIISLIDEGKQSITKLAAGWGSGSSRAY 51

```

```

RESULT 13
US-10-098-732A-33
; Sequence 33, Application US/10098732A
; Publication No. US20030175294A1
; GENERAL INFORMATION:
; APPLICANT: Skeiky, Yasir
; APPLICANT: Brannon, Mark
; APPLICANT: Guderian, Jeffrey
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Heterologous Fusion Protein Constructs Comprising a
; TITLE OF INVENTION: Leishmania Antigen
; FILE REFERENCE: 014058-012010US
; CURRENT APPLICATION NUMBER: US/10/098,732A
; CURRENT FILING DATE: 2003-04-29
; PRIOR APPLICATION NUMBER: US 60/275,837
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 33
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: ESAT-6
US-10-098-732A-33

```

Query Match	54.1%	Score 259;	DB 12;	Length 51;
Best Local Similarity	100.0%;	Pred. No. 4.2e-21;		

```

Matches      51; Conservative      0; Indels      0; Gaps      0;
OY          1 MTEQWNPAGTAAASATQGNVTGSIHSLDEGKOSLTFLAAAWGGSGSEAY 51
            |||||
DB          1 MTEQWNPAGTAAASATQGNVTGSIHSLDEGKOSLTFLAAAWGGSGSEAY 51
            |||||

RESULT 14
US-10-080-170-11
; Sequence 11, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495,0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; PRIORITY FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 11
; LENGTH: 95
; TYPE: PR1
; ORGANISM: Mycobacterium leprae
US-10-080-170-11
```

Query Match	38.8%	Score 186;	DB 16;	Length 95;
Best Local Similarity	36.3%;	Pred. No. 8.4e-13;		
Matches	33;	Conservative 27;	Mismatches 31;	Indels 0; Gaps 0;
QY	4	QQWNFAGIEAAGAIQGNVTSIHSILDEGKQSLTKLAAAWGGSGSEAYQGVQOKWDATAT	63	
Db	3	QAWHFPALQAVNELQGSQSRIDALLEQCOESLTKLQSSWNHGSNGSEYSSVQRRFNQTE	62	
QY	64	ELNNALQNLARTISEAGQAMASTEQNVTMF	94	
Db	63	GINHALGDLVQAINHSAETMQOETAGVWSMF	93	

```

RESULT 15
US-09-738-626-4145
; Sequence 4145, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738, 626
; CURRENT FILING DATE: 2000-12-18
; PRIOR APPLICATION NUMBER: JP 99/377484
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: JP 00/159162
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: JP 00/280988
; PRIOR FILING DATE: 2000-08-03
; NUMBER OF SEQ ID NOS: 7059
; SOFTWARE: PatentIn ver. 3.0
; SEQ ID NO 4145
; LENGTH: 96
; TYPE: FRT
; ORGANISM: Corynebacterium glutamicum

```

```

US-09-738-626-4145
Query Match      31.5%; Score 151; DB 10; Length 96;
Best Local Similarity 37.5%; Pred. No. 5.6e-09;
Matches 30; Conservative 14; Mismatches 36; Indels 0; Gaps 0;

Qy 5 QWNFAGIEAASAIQGNVTSIHSLDEGKQSLTKLAANGSGSEAYGVQVKWDATATE 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 RYEFCAIQGAATDINSTGRINSILDLGLKSLQLPWVASWEGESSEAYSEAQIKWDRAAAE 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 65 LNNALQNLARTISEAQAWA 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 LNTILATISNTVAQGAERMS 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 16
US-09-813-333-60
; Sequence 60, Application US/09813333
; Patent No. US20020119160A1
; GENERAL INFORMATION:
; APPLICANT: DeGroot, Anne S
; TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
; FILE REFERENCE: 17999-004 US
; CURRENT APPLICATION NUMBER: US/09/813,333
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,834
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-813-333-62
Query Match      20.5%; Score 98; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00048;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 IQGNVTSIHSLDEGKQSLT 37
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 IQGNVTSIHSLDEGKQSLT 20
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 19
US-10-044-703-62
; Sequence 62, Application US/10044703
; Publication No. US20020192233A1
; GENERAL INFORMATION:
; APPLICANT: DeGroot, Anne S
; TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
; FILE REFERENCE: 17999-004 US
; CURRENT APPLICATION NUMBER: US/10/044,703
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/190,834
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-044-703-62
Query Match      20.5%; Score 98; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00048;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 18 IQGNVTSIHSLDEGKQSLT 37
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 IQGNVTSIHSLDEGKQSLT 20
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 20
US-09-813-333-61
; Sequence 61, Application US/09813333
; Patent No. US20020119160A1
; GENERAL INFORMATION:
; APPLICANT: DeGroot, Anne S
; TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
; FILE REFERENCE: 17999-004 US
; CURRENT APPLICATION NUMBER: US/09/813,333
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,834
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61

US-09-738-626-4145
Query Match      31.5%; Score 151; DB 10; Length 96;
Best Local Similarity 37.5%; Pred. No. 5.6e-09;
Matches 30; Conservative 14; Mismatches 36; Indels 0; Gaps 0;

Qy 5 QWNFAGIEAASAIQGNVTSIHSLDEGKQSLTKLAANGSGSEAYGVQVKWDATATE 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 5 RYEFCAIQGAATDINSTGRINSILDLGLKSLQLPWVASWEGESSEAYSEAQIKWDRAAAE 64
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 65 LNNALQNLARTISEAQAWA 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 65 LNTILATISNTVAQGAERMS 84
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 16
US-09-813-333-60
; Sequence 60, Application US/09813333
; Patent No. US20020119160A1
; GENERAL INFORMATION:
; APPLICANT: DeGroot, Anne S
; TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
; FILE REFERENCE: 17999-004 US
; CURRENT APPLICATION NUMBER: US/09/813,333
; CURRENT FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,834
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-813-333-60
Query Match      21.1%; Score 101; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QWNFAGIEAASAIQGNVTS 24
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 QWNFAGIEAASAIQGNVTS 20
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 17
US-10-044-703-60
; Sequence 60, Application US/10044703
; Publication No. US20020192233A1
; GENERAL INFORMATION:
; APPLICANT: DeGroot, Anne S
; TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
; FILE REFERENCE: 17999-004 US
; CURRENT APPLICATION NUMBER: US/10/044,703
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/190,834
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-044-703-60
Query Match      21.1%; Score 101; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00023;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 QWNFAGIEAASAIQGNVTS 24
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 QWNFAGIEAASAIQGNVTS 20
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-813-333-61

Query Match      20.3%; Score 97; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 NFAGIEAAASAIQGNVTSIH 26
Db 1 NFAGIEAAASAIQGNVTSIH 20

RESULT 21
US-10-044-703-61
; Sequence 61, Application US/10044703
; Publication No. US20020192233A1
; GENERAL INFORMATION:
; APPLICANT: Degroot, Anne S
; TITLE OF INVENTION: Human T Cell Response to MHC-Binding Motif Clusters
; CURRENT APPLICATION NUMBER: US/10/044,703
; CURRENT FILING DATE: 2002-05-20
; PRIOR APPLICATION NUMBER: 60/190,834
; PRIOR FILING DATE: 2000-03-20
; NUMBER OF SEQ ID NOS: 81
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 61
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-044-703-61

Query Match      20.3%; Score 97; DB 14; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.00062;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 NFAGIEAAASAIQGNVTSIH 26
Db 1 NFAGIEAAASAIQGNVTSIH 20

RESULT 22
US-09-804-980-175
; Sequence 175, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 175
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-175

Query Match      20.0%; Score 96; DB 12; Length 96;
Best Local Similarity 30.2%; Pred. No. 0.0056;
Matches 29; Conservative 16; Mismatches 39; Indels 12; Gaps 3;

QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGSGSSEAYQGVQKQWDA 60
Db 1 MSQIMVNPAMAHAGDMAGYAGTQLSLGADIASEQAVLSSAWQDGTGITYQGMQTQW-- 58

RESULT 23
US-09-804-980-199
; Sequence 199, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 199
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-199

Query Match      20.0%; Score 96; DB 12; Length 96;
Best Local Similarity 30.2%; Pred. No. 0.0056;
Matches 29; Conservative 16; Mismatches 39; Indels 12; Gaps 3;

QY 1 MTEQOWNFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGSGSSEAYQGVQKQWDA 60
Db 1 MSQIMVNPAMAHAGDMAGYAGTQLSLGADIASEQAVLSSAWQDGTGITYQGMQTQW-- 58

RESULT 24
US-10-156-761-10069
; Sequence 10069, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10069
; LENGTH: 99
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10069

Query Match      18.6%; Score 89; DB 15; Length 99;
Best Local Similarity 22.4%; Pred. No. 0.034;
Matches 17; Conservative 19; Mismatches 40; Indels 0; Gaps 0;

QY 5 QNWFAGIEAAASAIQGNVTSIHSLDEGKQSLTKLAAAWGSGSSEAYQGVQKQWDA 64
Db 9 QVTYDALDEAATSIQNEAKTLEQDLQELKLVQCKQYWDGAAQDAFDSKLAWDKEAD 68

QY 65 LNNALQNLTARTISEAG 80
Db 69 IHTALTGTGHAAGTAG 84
```

Db 1 MSQIMYNYPAMLGHAGDMAGYAGTQLSLGABIAVEQAALQSAWQDGTGITYQAWQAQW-- 58
Qy 61 TATELNALQNLARTISEAGQAMAST-EGNVVTGMFA 95
Db 59 -----NOAMEDLVR-----AYHAMSGSTHEANTMAMMA 85
RESULT 27
US-09-804-980-195
; Sequence 195, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 195
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-195
Query Match 17.7%; Score 85; DB 12; Length 96;
Best Local Similarity 29.2%; Pred. No. 0.089;
Matches 28; Conservative 14; Mismatches 42; Indels 12; Gaps 3;
Qy 1 MTEQQWNFAGIEAAASAIQGNVTSHSLLDGKQSLTKLAAAWGSGSEAYQGVQKQWDA 60
Db 1 MSQIMYNYPAMLGHAGDMAGYAGTQLSLGABIAVEQAALQSAWQDGTGITYQAWQAQW-- 58
Qy 61 TATELNALQNLARTISEAGQAMAST-EGNVVTGMFA 95
Db 59 -----NOAMEDLVR-----AYHAMSGSTHEANTMAMMA 85
RESULT 28
US-10-080-170-378
; Sequence 378, Application US/10080170
; Publication No. US20030129601A1
; GENERAL INFORMATION:
; APPLICANT: COLE, S.T.
; TITLE OF INVENTION: COMPARATIVE MYCOBACTERIAL GENOMICS AS A TOOL FOR
; TITLE OF INVENTION: IDENTIFYING TARGETS FOR THE DIAGNOSIS, PROPHYLAXIS OR
; TITLE OF INVENTION: TREATMENT OF MYCOBACTERIOSES
; FILE REFERENCE: 03495.0218
; CURRENT APPLICATION NUMBER: US/10/080,170
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: 60/270,123
; PRIOR FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 652
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 378
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-10-080-170-378
Query Match 17.7%; Score 85; DB 16; Length 96;
Best Local Similarity 29.2%; Pred. No. 0.089;
Matches 28; Conservative 14; Mismatches 42; Indels 12; Gaps 3;
Qy 1 MTEQQWNFAGIEAAASAIQGNVTSHSLLDGKQSLTKLAAAWGSGSEAYQGVQKQWDA 60
Db 1 MSQIMYNYPAMLGHAGDMAGYAGTQLSLGABIAVEQAALQSAWQDGTGITYQAWQAQW-- 58
Qy 61 TATELNALQNLARTISEAGQAMAST-EGNVVTGMFA 95
Db 59 -----NOAMEDLVR-----AYHAMSGSTHEANTMAMMA 85

US-09-791-171-2
; Sequence 2, Application US/09791171
; Patent No. US20020094336A1
; GENERAL INFORMATION:
; APPLICANT: ANDERSEN, Peter
; APPLICANT: NIELSEN, Rikke
; APPLICANT: OETTINGER, Thomas
; APPLICANT: RASMUSSEN, Peter Birk
; APPLICANT: ROSENKEANDS, Ida
; APPLICANT: WELDINGH, Karin
; APPLICANT: FLORIO, Walter
; TITLE OF INVENTION: NUCLEIC ACIDS FRAGMENTS AND POLYPEPTIDE FRAGMENTS
; TITLE OF INVENTION: DERIVED FROM M. TUBERCULOSIS
; FILE REFERENCE: 670001-2002.1
; CURRENT APPLICATION NUMBER: US/09/791,171
; CURRENT FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 09/050,739
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 0376/97
; PRIOR FILING DATE: 1997-04-02
; PRIOR APPLICATION NUMBER: 1277/97
; PRIOR FILING DATE: 1997-11-10
; PRIOR APPLICATION NUMBER: 60/044,624
; PRIOR FILING DATE: 1997-04-18
; PRIOR APPLICATION NUMBER: 60/070,488
; PRIOR FILING DATE: 1998-01-05
; NUMBER OF SEQ ID NOS: 173
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-791-171-2
Query Match 17.7%; Score 85; DB 9; Length 96;
Best Local Similarity 29.2%; Pred. No. 0.089;
Matches 28; Conservative 14; Mismatches 42; Indels 12; Gaps 3;
Qy 1 MTEQQWNFAGIEAAASAIQGNVTSHSLLDGKQSLTKLAAAWGSGSEAYQGVQKQWDA 60
Db 1 MSQIMYNYPAMLGHAGDMAGYAGTQLSLGABIAVEQAALQSAWQDGTGITYQAWQAQW-- 58
Qy 61 TATELNALQNLARTISEAGQAMAST-EGNVVTGMFA 95
Db 59 -----NOAMEDLVR-----AYHAMSGSTHEANTMAMMA 85
RESULT 26
US-09-804-980-2
; Sequence 2, Application US/09804980
; Publication No. US20030147897A1
; GENERAL INFORMATION:
; APPLICANT: Statens Serum Institut
; APPLICANT: Anderson, Peter
; TITLE OF INVENTION: M. Tuberculosis Antigens
; FILE REFERENCE: 670001-2002.4
; CURRENT APPLICATION NUMBER: US/09/804,980
; CURRENT FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 257
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 96
; TYPE: PRT
; ORGANISM: Mycobacterium tuberculosis
US-09-804-980-2
Query Match 17.7%; Score 85; DB 12; Length 96;
Best Local Similarity 29.2%; Pred. No. 0.089;
Matches 28; Conservative 14; Mismatches 42; Indels 12; Gaps 3;
Qy 1 MTEQQWNFAGIEAAASAIQGNVTSHSLLDGKQSLTKLAAAWGSGSEAYQGVQKQWDA 60

```

RESULT 30
US-09-916-201-7
; Sequence 7, Application US/09916201
; Patent No. US20020131976A1
; GENERAL INFORMATION:
; APPLICANT: LALVANT, Ajit
; APPLICANT: PATHAN, Ansar A.
; APPLICANT: HILL, Adrian V.S.
; TITLE OF INVENTION: TUBERCULOSIS VACCINE
; FILE REFERENCE: 117-359
; CURRENT APPLICATION NUMBER: US/09/916,201
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 09/467,893
; PRIOR FILING DATE: 1999-12-21
; PRIOR APPLICATION NUMBER: US 60/113,783
; PRIOR FILING DATE: 1998-12-23
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 7
; LENGTH: 15
; TYPE: PRT
; ORGANISM: M.tuberculosis
US-09-916-201-7

```

Query Match 16.9%; Score 81; DB 10; Length 15;

```

Best Local Similarity 100.0%; Pred. No. 0.024;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      46  SGSEAYQGVQOKWDA  60
      |||||
Db       1  SGSEAYQGVQOKWDA  15

Search completed: February 5, 2004, 17:43:18
Job time : 15.119 secs

```

